PROCEEDINGS OF THE TWENTY-NINTH ANNUAL MEETING OF THE

BERKELEY LINGUISTICS SOCIETY

February 14-17, 2003

GENERAL SESSION

and

PARASESSION on PHONETIC SOURCES OF PHONOLOGICAL PATTERNS: SYNCHRONIC AND DIACHRONIC EXPLANATIONS

Edited by

Pawel M. Nowak
Corey Yoquelet
David Mortensen

Berkeley Linguistics Society
Berkeley, California, USA

PROCEEDINGS OF THE TWENTY-NINTH ANNUAL MEETING OF THE

BERKELEY LINGUISTICS SOCIETY

February 14-17, 2003

GENERAL SESSION

and

PARASESSION on PHONETIC SOURCES OF PHONOLOGICAL PATTERNS: SYNCHRONIC AND DIACHRONIC EXPLANATIONS

Edited by

Pawel M. Nowak
Corey Yoquelet
David Mortensen

Berkeley Linguistics Society
Berkeley, California, USA

Berkeley Linguistics Society
University of California, Berkeley
Linguistics Department
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA

ISSN 0363-2946

LCCN 76-64-0143

Printed by: Sheridan Books, Inc.
P.O. Box 1203
Ann Arbor, Michigan 48108
USA

TABLE OF CONTENTS

GENERAL SESSION

Differential Coding, Partial Blocking and Bidirectional OT — 1
JUDITH AISSEN

Does *Most* Mean 'More Than Half'? — 17
MIRA ARIEL

Functional Specialization in Tone Perception: Evidence from Dichotic Listening in Yalalag Zapotec — 31
HERIBERTO AVELINO

Variability and Constancy in the Articulation and Acoustics of Pima Coronals — 43
HERIBERTO AVELINO and SAHYANG KIM

Another Look at Australia as a Linguistic Area — 55
CLAIRE BOWERN

Quantity, Causality and Temporality in Change Constructions — 67
CRISTIANO BROCCIAS

Preemptive Usage of N1 *to yuu* N2 in Japanese Conversation: Psychological Strategy to Increase the Distance between the Hearer and the Entity — 79
MICHICO TODOKORO BUCHANAN

WANT TO/WANNA: Verbal Polysemy vs. Constructional Compositionality — 91
GUILLAUME DESAGULIER

Phonological Contrasts and Phonetics in Manchu Vowel Systems — 103
B. ELAN DRESHER and XI ZHANH

I Bet you Think this Paper is about *you*: Participant Roles and *you* — 115
RANDALL EGGERT

Shoshoni Verb Classes and the Perception of Aspiration — 127
DIRK ELZINGA

Syntactic and Semantic Valence: Morphosyntactic Evidence from Minangkabau — 137
CATHERINE FORTIN

A Study of the Speaking Fundamental Frequency Characteristics and Perceived Pitch Characteristics of Black and White Women — 149
NICOLE GASKINS

The Status of the Morpheme in Georgian Verbal Morphology — 161
OLYA GUREVICH

Definition of the Base — 173
ANYA HOGOBOOM

Lexical Representation in Korean: Evidence from the Tip of the Tongue Phenomenon — 185
SUNYOUNG HONG

Markedness and the Phonetic Implementation of Tone in North Kyungsang Korean — 197
WILLIAM IDSARDI and WOOHYEOK CHANG

Multiple Questions in Basque — 209
YOUNGMI JEONG

Nominal and Adjectival Roots in North Saami Verbs — 221
MARIT JULIEN

Syllabically Conditioned Perceptual Epenthesis — 233
BARIS KABAK and WILLIAM IDSARDI

The Meanings of Consonants — 245
BARBARA F. KELLY, WILLIAM R. LEBEN and ROBERT H. COHEN

Morphologically Motivated Prosodic and Metrical Structures — 255
EUN-SOOK KIM

English Locative Inversion: Grammatical Interfaces and Constructions — 267
JONG-BOK KIM

Contrast and Phonological Activity in the Nez Perce Vowel System | 283
SARA MACKENZIE and B.ELAN DRESHER

On Morphological Case and Word-Order Freedom | 295
THOMAS MCFADDEN

Subsumption in Feature Theory and Speech Recognition | 307
MORITZ NEUGEBAUER

A Bipartite Verb Stem Outlier in Eurasia: Nakh-Daghestanian | 321
JOHANNA NICHOLS

Establishing a Community of Practice on the Internet: Linguistic Behavior in Japanese Online Communication | 337
YUKIKO NISHIMURA

Rethinking 'Thinking For Speaking' | 349
STEPHANIE POURCEL

The Japanese 'Nominative' Particle: A New Approach | 359
BENJAMIN ROSENTHAL

Lexical Confusability and Degree of Coarticulation | 367
REBECCA ANNE SCARBOROUGH

On the Syntactic Status of Certain Fronted Adverbials in English | 379
BENJAMIN SHAER

Patterns of Semantic Harmonization in English: The Case of May Well | 391
REIJIROU SHIBASAKI

Constraints on Synaesthesia | 403
YOSHIKATA SHIBUYA and HAJIME NOZAWA

Reexamining Condition C Effects | 415
KAYONO SHIOBARA

Analogical Modeling: Exemplars, Rules, and Quantum Computing | 425
ROYAL SKOUSEN

A Common Structure for Cross-Linguistics Conjunction Patterns | 441
MARCUS SMITH

Indexicals in Dialogic Interaction 451
MAKIKO TAKEKURO

The Causatives in Sun Hongkai's Anong: 463
Language Death and Rapid Restructuring
GRAHAM THURGOOD

Junctural and Parasitic Voicing in Burmese 473
KENNETH VANBIK

PARASESSION: PHONETIC SOURCES OF PHONOLOGICAL PATTERNS – SYNCHRONIC AND DIACHRONIC EXPLANATIONS

Phonetics and Phonology of Transparent Vowels in Hungarian 485
STEFAN BENUS, ADIAMANTIS GAFOS
and LOUIS GOLDSTEIN

Antigemination: Natural or Unnatural History? 499
JULIETTE BLEVINS

Children's Unnatural Phonology 523
EUGENE BUCKLEY

Consonant Confusability: An MEG Study 535
VALENTINE HACQUARD and MARYANN WALTER

Positional Markedness as a By-prouct of the Learning Situation: 545
a Simulation
ALEXEI KOCHETOV

The Diachronic Influence of Perception: 557
Experimental Evidence from Turkish
JEFF MIELKE

Accepting Markednesslessness: How Non-phonological Symbolic 569
Computation Shapes Trends in Attested Phonological Systems
CHARLES REISS

Knowledge of Similarity and Narrow Lexical Override 583
DONCA STERIADE

The Non-neutralizing Nature of Hungarian Voicing Assimilation 599
ZOË TOFT and WOUTER JANSEN

Self-organization and Categorical Behavior in Phonology 611
ANDREW WEDEL

Some Methodological Issues in Doing Phonetic Typology: Cantonese Contour Tone Revisited 623
ALAN C. L. YU

General Session

Differential Coding, Partial Blocking, and Bidirectional OT*

JUDITH AISSEN
UC Santa Cruz

1. Differential coding and partial blocking

PARTIAL BLOCKING – in which two semantically equivalent forms coexist, but with non-overlapping distributions – poses a problem for Optimality Theory. A classic OT evaluation produces one winner, but in partial blocking there are (at least) two. Recent work has proposed WEAK BIDIRECTIONAL OT as one solution to partial blocking (Blutner 1999), but the range of phenomena which motivate this architecture over alternatives remains relatively unexplored. Here I suggest that certain cases of DIFFERENTIAL CODING provide one motivation for Weak bidirectional OT.

I use the term DIFFERENTIAL CODING to refer to morphology or syntax which (a) codes the grammatical function of the core arguments in transitive or passive clauses, and (b) is differential in the sense that it is selective, with its distribution dependent on semantic and pragmatic features of the arguments. Differential coding may involve case marking, agreement, direction, or voice. Well-known examples are subject case marking in many Australian languages (e.g. Dyirbal), where some transitive subjects are case marked but not all, differential object marking (e.g. Spanish, Hindi) where some direct objects are case marked but not all. It includes inverse marking as in the Algonquian languages or Nocte, where the verb may carry a mark depending on semantic and pragmatic properties of both nominal arguments (DeLancey 1981). And in some languages, e.g. Lummi, voice is differential in that the distribution of active and passive is categorically restricted by semantic and pragmatic features of agent and patient (Jelinek and Demers 1983).

There is a relation between differential coding and markedness (or prototypicality). Certain semantic and pragmatic properties are prototypical (statistically more frequent) for grammatical objects, especially low animacy, low definiteness, and low topicality. Transitive subjects, in contrast, are prototypically high in animacy, definiteness, and topicality. Differential coding systems mark

* This research was supported by the National Science Foundation, Grant #SBR-9818177. Thanks also to Joan Bresnan, Gerhard Jäger, and Arnold Zwicky, whose input has influenced this work.

subjects and/or objects which diverge from the prototype, leaving unmarked those which are more prototypical, a generalization known from typological work:

> ...the most natural kind of transitive construction is one where the A (=transitive subject, JA) is high in animacy and definiteness, and the P (=transitive object, JA) is lower in animacy and definiteness; and any deviation from this pattern leads to a more marked construction. (Comrie 1989:128)

Differential coding is simply the kind of more marked construction which marks deviations from the norm.

Some of these norms are listed in (1). The prototypical or unmarked situation is for the transitive subject (henceforth 'S') to be associated with the high end of the dimensions in (1a-e) – to be 1st or 2nd person (i.e. LOCAL PERSON), to be animate, to be definite, to be topical. Conversely, the unmarked situation for the direct object ('O') is to be associated with the low end of the same dimensions – to be 3rd person, inanimate, indefinite, non-topic.

(1) Unmarked alignments

a. Gram. Function:	S	>	O
	↓		↓
b. Person:	1st/2nd		3rd
c. Animacy:	Human		Inanimate
d. Definiteness:	Definite		Indefinite
e. Topicality:	High		Low

In Dyirbal, it is 3rd person S's (marked S's) which are case marked, not 1st or 2nd; it Spanish, it is animate and specific O's (marked O's) which are case marked, not inanimate or non-specific ones; in Nocte, it is clauses in which both S and O are marked (3rd person S, local person O), which get inverse marking; and in Lummi, it is active clauses with 3rd person S and local person O which are blocked, requiring instead passive expression. It is possible then to derive from (1) a set of implicational universals regarding the distribution of differential coding (on this, see, among others, (Silverstein 1976, Comrie 1989, Aissen 1999a, 2003).

2. The pragmatic division of labor

Differential coding is an instance of PARTIAL BLOCKING (cf. Horn 1989:194). Partial blocking refers to cases in which two semantically equivalent forms coexist, but have non-overlapping distributions. One form, generally the less complex one, expresses a core or prototypical sense, while the other, usually a more complex form, is relegated to a more peripheral sense. An example from Kiparsky (1982) is the pair *cook/cooker*. *Cook* is an unmarked agentive noun expressing what we would expect *cooker* to mean; *cooker* is not blocked by *cook*, but it has a more peripheral sense referring to a utensil.[1] Another example, from McCawley (1978), is the pair *kill/cause to die.* The lexical causative *kill* usually describes direct causation, the periphrastic *cause to die* indirect causation. In McCawley's pragmatic account, *cause to die* is restricted to indirect causation because the expression of direct causation is preempted by a simpler form. *Kill* doesn't entirely block *cause to die*, but it blocks one sense.

More generally, McCawley and Horn suggest that partial blocking involves a division of pragmatic labor whereby the less marked (less complex) form expresses the less marked meaning – the one associated with the more stereotypical situation – with the more marked (more complex) form relegated to covering the more marked meaning. McCawley and Horn appeal to Gricean maxims of quantity to explain the association between markedness of form and the stereotypicality of situation. The first is a speaker-oriented economy condition: *don't make your contribution more informative than is required.* If you intend to refer to a prototypical situation, you can afford to be brief. The second is hearer-oriented: *make your contribution as informative as is required.* If you intend to refer to a situation which is not prototypical, you may need to say more because if you use the economical form, the hearer will assume you are referring to the prototypical situation.

3. Differential coding and optimization

There is a clear connection between differential coding and the Gricean account of partial blocking sketched above. When associations are prototypical or unmarked, they need not be formally marked. When they are not prototypical, special marking is in order. The functional motivation of differential coding thus seems clear – it has to do with recoverability of grammatical function. This invites us to ask whether recoverability actually plays a role in determining the synchronic distribution of differential coding. In many cases, perhaps most, the answer is *no.* Differential coding is often overextended (generalized) and obligatory even where it is not needed to recover grammatical function. In Spanish, *Veo *(a) Juan,* 'I see Juan', the case-marking preposition *a* is required for *Juan* even though verb agreement (1st singular subject) leaves no doubt that *Juan* is the direct object. Similarly, Dyirbal has both differential subject marking AND differential object marking, but the two operate independently (Dixon 1972). As a result, there are clauses which require overt case on BOTH S and O (e.g. a clause with 3rd person S, 1st person O). Case on one or the other would be sufficient to insure recoverability of grammatical function; we don't need both.

In short, while there are implicational universals in the distribution of differential coding, it is still necessary in many cases to stipulate exactly what gets marked and what doesn't. The fact that definite inanimate objects in Hindi may be

[1] Partial blocking contrasts with (full) blocking, "the non-occurrence of one form…due to the simple existence of another" (Aronoff 1976:43). E.g. *men* blocks **mans,* and *fury* blocks **furiosity.*

case marked, while in Spanish they cannot be, probably does not follow from other more general properties of these grammars. Within an OT context, 'stipulate' means that the distribution of case marking is enforced by a quite specific constraint, as in (2):

(2) MARK OBJ/HUM & SPEC » *CASE » MARK OBJ

Here, the top constraint enforces case marking for human, specific objects and outranks a constraint penalizing case marking (*CASE). Case marking of other objects is turned off because *CASE outranks, MARK OBJ, a general contraint enforcing case marking of objects (cf. Aissen 2003).

However, there are other cases of differential coding where the distribution of the marked element, while also highly restricted, seems to be exactly determined by recoverability, i.e. overt marking is required exactly when its absence would lead to an interpretation other than the one intended. One example is the distribution of the 3rd person object marker in Takelma; another is the agent focus verb form in Tzotzil. In both cases, the distribution of differential coding is parasitic on the interpretation of the unmarked form and therefore should not be stipulated. Ideally, the analysis would stipulate little about the morphemes involved, allowing their distribution to follow from the way the unmarked form is interpreted. To achieve this, it is necessary to optimize both over form and interpretation.

4. Takelma

Takelma is no longer spoken. My discussion is based on Culy (2000), which was in turn based on Sapir (1922). The morphology of interest here is the suffix *-k^h^wa*, a 3rd person object marker (OM) which is realized on the verb. The distribution of *-k^h^wa* is very restricted, but I suggest that it is NOT determined by a highly specific constraint like the top one in (2), but by the interaction of more general constraints.

Takelma has a full set of object markers, shown in (3).

(3)

Object markers	Singular	Plural
1st	-xi	-am
2nd	-pi	-anp^h
3rd	ø /-k^{h}wa	ø /-k^{h}wa

Overt marking is required for 1st and 2nd person objects, with singular and plural distinguished. (Overt agreement is also required for 1st and 2nd person subjects, but is not shown here.) For 3rd person objects, there is an alternation between ø (zero) and the suffix *–k*^{h}*wa*; there is no distinction between singular and plural.

Culy documents that the ø variant is found with O's of all kinds, including animates (4a) and inanimates (4b), covert pronouns (4c) and overt nominals (4a). The zero variant occurs in combination with subjects of all three persons (4a, b, c).[2]

(4) a. ani:ʔ ki: t'omoàʔn hamìʔt^{h}pan. (human object (TT 158:3))
NEG 1SG kill-1SG father-2PL
'I did not kill your father.'

b. k^{h}ài nakai:t^h? (inanimate object (TT 56:9))
something do/say-2SG
'What did you say?'

c. alsinlò:k^h. (covert object (TT 24:12))
meet
'They met him.'

The overt variant *-k*^{h}*wa*, on the other hand, is subject to three constraints. First, *-k*^{h}*wa* only occurs when the S is also 3rd person (Sapir 1922:168). Second, following Culy, *-k*^{h}*wa* is always used when O is higher in animacy than S. Usually, this means that the object refers to a human, as in (5b), or to a "mythic animal conceived of as a human being" (Sapir 1922:168), (5a).

(5) a. mi: p'owó:k^{h}wa
now sting-OM
'Now they (the yellowjackets) stung him (Coyote).' (TT 74:3)

b. mená yap'a t'omò:k^{h}wa
bear man kill-OM
'The bear killed the man.' (TG 158)

Third, when S and O are of equal animacy, *-k*^{h}*wa* occurs when O is topical, i.e. when it outranks S in topicality. In such cases, the O is usually pronominal (and covert). In both (6a, b), the O is topical at the point in the text where the sentence occurs.

[2] Most of the Takelma examples are taken directly from Culy 2000 (in the IPA orthography that he uses), but are identified according to the original source. Text examples (*TT*) are from Sapir's *Takelma Texts*, reprinted in Golla (1990). Other examples (*TG)* are from Sapir 1922, also reprinted in Golla (1990). Page numbers refer to those of the original publications.

(6) a. há:xan<u>k^{h}wa</u>hì:s.
burn-OM-almost
'He [Sinew-man] almost burned him [Daldal].' (TT 27:16)

b. xamkwitìk^{hw}ta<u>kwa</u> mé:x.
threw.into.water-OM crane
'Crane threw her [Grizzly bear] into the water.' (TT 122:13)

Sapir saw the distribution of *-k^{h}wa* in terms of recoverability, and cited the pair in (7) to illustrate (1922:169).

(7) a. t'ipisi: t'ayák^h.
ants found
'He found the ants.'

b. t'ipisi: t'ayá:<u>k^{h}wa</u>.
ants found-OM
'The ants found him.'

He observed that (7a) has only the interpretation 'he found the ants'. If, as in (7b), the intended interpretation were, 'the ants found him', where the O refers to a human and the S is overt, then *-k^{h}wa* is needed. Otherwise, the ABSENCE of *-k^{h}wa*, as in (7a), leads to the inverse interpretation. It is the presence of *–k^{h}wa* which somehow deflects that interpretation and forces the overt nominal to be parsed as subject.[3]

The account I develop below builds on this conception of *-k^{h}wa* and expresses the generalization in (8).

(8) *-k^{h}wa* occurs only when its absence would make unrecoverable the intended linking of nominal arguments to grammatical function.

The three restrictions on *-k^{h}wa* identified above can be (informally) understood in the following terms: *-k^{h}wa* is limited to clauses with 3rd person S and 3rd person O because with any other combination of persons, obligatory agreement with 1st and 2nd person S and O unambiguously determines the linkage of nominal arguments to grammatical function. In clauses then with one or two local person arguments,
-k^{h}wa is not needed to recover the intended interpretation, and in accord with (8), it does not occur. However, in clauses WITH 3rd person S and O, linkage to grammatical function is underdetermined by the agreement morphology; in some of those cases, but not all, *-k^{h}wa* is needed to determine that linkage.

When is it needed? Basically in clauses that run counter to (some of) the biases mentioned earlier – in particular, the bias that subjects be human and the bias that they be topical. Clauses which accord with those biases are unmarked

[3] See also fn. 11.

and require no special coding. This is the case in Sapir's (7a), where the subject is human and topical, and where -k^h*wa* is not used. However, clauses not in accord with these biases are marked and require the additional coding that -k^h*wa* provides (e.g. 5, 6, 7b). Agreement and bias play a crucial role then in delimiting the distribution of -k^h*wa* and this role should be reflected in the analysis.

5. (Weak) bidirectional optimization

Classic OT syntax took the perspective of production, in the sense that the input to evaluation was a meaning representation, and the candidate set consisted of alternative expressions of that meaning (Grimshaw 1997). Conversely, Hendriks and de Hoop (2001) have argued that OT semantics takes the perspective of comprehension, in the sense that the input to evaluation is a surface expression, with the candidate set consisting of alternative interpretations of that expression. However, various empirical problems, including partial blocking, suggest that optimization requires both perspectives at once and further, that this 'bidirectional' optimization proceeds under a single grammar (i.e. a single constraint ranking) (Blutner 1999).

To develop a bidirectional analysis of Takelma -k^h*wa*, we need to make explicit the relevant constraints, as well as their ranking. First, I assume that higher ranked constraints than the ones I discuss here eliminate various structures NOT found in Takelma (e.g. case marking, passives with overt agents). What IS relevant here are constraints which enforce agreement with local (1^{st}, 2^{nd}) person subjects and objects (9a), and one which enforces agreement with 3^{rd} person objects (9b). I refer to this simply as K^HWA below. In line with the strategy sketched at the end of §3, the constraint in (9b) is simple and makes no reference to any of the three restrictions on –k^h*wa* documented in §4. In tension with constraints enforcing agreement is one which penalizes agreement, *AGR (9c). *AGR is violated once for each nominal argument linked to verb agreement.

(9) a. AGR/OBJECT$_{LOCAL}$, AGR/SUBJECT$_{LOCAL}$
   b. AGR/OBJECT$_{3RD}$ (K^HWA)
   c. *AGR
   d. Bias: SUBJECT HUMAN, SUBJECT TOPIC

Finally, I assume that the biases discussed earlier function in grammars as markedness constraints. The relevant constraints are SUBJECT HUMAN and SUBJECT TOPIC (9d). These are satisfied if the subject is human and topic (respectively), but are violated otherwise. [4]

A little reflection makes clear that a classic unidirectional OT model based on the constraints in (9) cannot derive the distribution of ø and -k^h*wa*, for there is no

[4] For a more general discussion of Bias constraints, see Aissen (1999a, 2003).

ranking of *AGR and K^{H}WA that works. If *AGR » K^{H}WA, *-k^{h}wa* will never surface; if K^{H}WA » *AGR, ø will never surface. However, the right results can be achieved in a bidirectional model under the ranking *AGR » K^{H}WA. The overall ranking is shown in (10). We will build up to the crucial case, starting with simpler cases in which the ranking of (10) yields the right results in a classic, unidirectional (production-oriented) model.

(10) AGREEMENT: AGR/OBJECT$_{LOCAL}$ — AGR/SUBJECT$_{LOCAL}$

*AGR

AGR/OBJECT$_{3}$ (K^{H}WA)

BIAS: SUBJECT HUMAN

SUBJECT TOPIC

Consider first inputs with a local person S or O, as in (11). The input meaning, *I found him,* is shown at the bottom of the tableau. There are three candidate forms, which differ in how much agreement they carry. The most economical form, *f1*, has no agreement so violates the high-ranking constraint that enforces agreement with a 1st person S. The least economical form is *f3*, which agrees with both S and O. It violates *AGR twice, once more than the intermediate form, *f2*, which is thus correctly determined as output. Bias constraints play no role here.

(11)

Context: i = Topic	**AGR/ SUBJECT$_{1ST}$**	***AGR**	***K^{h}WA***	**SU TOP**
f1 found	*!		*	*
+*f2 found-1SG$_{SUBJ}$*		*	*	*
f3 found- 3$_{OBJ}$ - 1SG$_{SUBJ}$		*!*		*
	m: I found him$_{i}$.			

The tableau in (12) shows an evaluation for a clause with 3rd person S and O, with the meaning *he found the ants*. The referent of *he* is the topic (cf. 7a). The subject is human and topical, the object non-topical and non-human. In this case, there are two relevant candidates, one with the object marker (*f1*) and one without (*f2*). In addition to the surface string and the morphological analysis, these candidates are associated with a syntactic structure (not shown) which links the subject with a null pronoun and the object with the nominal 'ants'. The input also includes enough information about context to identify the topic (shown in the upper left corner).

(12)

Context: i = Topic	*AGR	K^hWA	SU HUM	SU TOP
+*f1 t'ipisi: t'ayák^h* ants found		*		
f2 t'ipisi: t'ayák^{h}wa *ants found-3$_{OBJ}$*	*!			
	▯1: he$_i$ found the ants			

Both candidates satisfy the Bias constraints since the subject is in each case human, and topical. They differ on the other two constraints. Under the proposed constraint ranking, the more economical *f1* wins, correctly blocking *f2*.

The difficult case for the ranking in (10) is the one in which *-k^{h}wa* DOES surface. This is an input with 3rd person S and O, where O outranks S in animacy or, if they are both animate, where O is topical, e.g. one with the meaning *the ants found him*, as in (13) (cf. 7b). The two candidates have the same surface form as those in (12) – one with the object marker (*f2*) and one without (*f1*). The syntax (not shown) associates the subject with 'ants' and the object with the null pronoun. In this case, both candidates VIOLATE the two Bias constraints since in neither case is the subject human or topical.

(13)

Context: i = Topic	*AGR	K^hWA	SU HUM	SU TOP
7f1 t'ipisi: t'ayák^h ants found		*	*	*
f2 t'ipisi: t'ayák^{h}wa *ants found-3$_{OBJ}$*	*!		*	*
	▯2: the ants found him$_i$			

The problem is that *f1* is STILL optimal since it is the more economical form. This is the wrong result (indicated by *7*) since this meaning is in fact expressed by *f2*, not *f1*. The fact that *f2* has a worse constraint profile than *f1* cannot be solved by reranking the two top constraints. That would force agreement with ALL 3rd person objects, yielding the right result for (13), but the wrong result for (11) and (12).

The key intuition here is that *f1* cannot express *m2* because it is surface identical to *f1* in (12), and *m1* is a better interpretation for *f1* than *m2* is. Thus it is the pairing of *f1* with *m1* in (12) which blocks the pairing of *f1* with *m2* in (13). To achieve this formally, we must take into account not only the optimal expression of *m2*, but also the optimal interpretation of (the unparsed surface string associated with) *f1*.

Putting (12) and (13) side-by-side makes clearer what is needed. Pairing the two forms {*f1, f2*} with the two meanings {*m1, m2*} yields four *<form, meaning>* pairs, represented by the four quadrants in (14). The Eval function should output two optimal pairs: *<f1, m1>* and *<f2, m2>* – the unmarked form is paired with the unmarked meaning (see +), the marked form with the marked meaning (see ʃ).

(14)

Context: i = Topic	*AGR	K^hWA	SU HUM	SU TOP	*AGR	K^hWA	SU HUM	SU TOP
+*f1 t'ipisi: t'ayák*^{h} ants found		*						
ʃ*f2 t'ipisi: t'ayák*^{h}*wa* *ants found-3*$_{OBJ}$					*		*	*
	+ □1: he$_i$ found the ants				ʃ □2: the ants found him$_i$.			

A number of bidirectional OT architectures have been proposed as solutions to particular empirical problems (see Beaver and Lee (to appear) for a survey and discussion). Several of these cannot account for the partial blocking which characterizes Takelma -*k*^{h}*wa/ø*. One is STRONG OT OPTIMALITY (Blutner 1999). Beaver and Lee (to appear) define it as follows: a form-meaning pair *<f, m>* is *Strong OT optimal* iff (a) *<f, m>* ∈ GEN; (b) there is no pair *<f', m>* such that *<f', m>* ≻ *<f, m>*;[5] and (c) there is no pair *<f, m'>* such that *<f, m'>* ≻ *<f, m>*. There is only one Strong OT optimal pair in (14), which is *<f1, m1>*. Hence Strong OT optimality is too strong to account for Takelma, where there are two optimal pairs. This is a general property of Strong OT optimality, one which motivates WEAK OT OPTIMALITY, also introduced in Blutner (1999).

Weak OT optimality is similar to Strong OT optimality, but is recursive.[6] Having identified *<f1, m1>* in (14) as optimal, suboptimal pairs which share with it either a form or a meaning are eliminated from further competition. In (14), these are the shaded quadrants, *<f1, m2>* and *<f2, m1>*. With these candidates eliminated, we search for new Weak OT optimal pairs. Although *<f2, m2>* has the worst profile of all in (14), it is a Weak OT optimal pair since there is now no Weak OT optimal pair which shares either a form or a meaning with *<f2, m2>* and which is more harmonic than it. In short, there is no better interpretation for *f2* than *m2*, and no better expression of *m2* than *f2*.

---

[5] Read '≻' as 'more harmonic than', where relative harmony is defined as in classic OT.

[6] Beaver and Lee provide this definition for WEAK OT OPTIMAL: *<f, m>* is *Weak OT optimal* iff (a) *<f, m>* ∈ GEN; (b) there is no Weak OT optimal *<f', m>* ∈ GEN such that *<f', m>* ≻ *<f, m>*; and (c) there is no Weak OT optimal *<f, m'>* ∈ GEN such that *<f, m'>* ≻ *<f, m>*. Clauses (b) and (c) in this definition correspond to the Gricean principles mentioned earlier (see Blutner (1999) for discussion).

The Takelma facts provide empirical support for Weak OT optimality over the Strong version.[7] Weak OT optimality derives the distribution of *-k^h^wa*, and it does so through the interaction of general constraints. Bias constraints and constraints governing agreement with *local* person S and O bear the brunt of the analytical burden, rather than a complex constraint on *-k^h^wa* itself.

6. Tzotzil[8]

The relation between the two 3rd person object markers in Takelma is not unique. Tzotzil has a bit of morphosyntax whose distribution is remarkably similar to that of Takelma *-k^h^wa*. This is the so-called AGENT FOCUS verb form which occurs only when subject of a transitive clause is extracted, as in WH questions, relative clauses, and focus (Aissen 1999b). The AF form is derived from a transitive verb stem by suffixing *–on*: [9]

(15) a. Pero buch'u i-mil-on?
but who CP-kill-AF
'But who killed her?'

b. K'usi i-sibtas-on li antzetike?
what CP-frighten-AF the women
'What frightened the women?'

Aside from the restriction to subject extraction contexts, the AF form occurs in the same contexts as Takelma *-k^h^wa*. It occurs only in clauses in which both S and O are third person, and it only occurs when the O > S in animacy or topicality (Aissen 1999b). The examples in (15) are typical: in (a) O > S in topicality; in (b), O > S in animacy. In all other contexts, the plain transitive verb occurs. I.e. it occurs when one or both of the core arguments is a local person; and it occurs when S is outranked by O in neither animacy nor topicality. In both examples in (16), the S is animate and the O is inanimate, and in neither case is the object topical.

(16) a. Buch'u i-s-pas mantal?
who CP-E3-make order
'Who's giving the orders?'
b. li vinik [ta x-chon paxak] e...
the man ICP E3-sell pineapple-ENC

---

[7] This is not to say that there are no problems facing weak OT optimality. See Beaver and Lee (to appear) for discussion. Another variant of bidirectional OT which cannot account for the Takelma facts is that of Wilson (2001). Space limitations preclude discussion here.

[8] Stiebels (2003) has independently suggested the desirability of a bidirectional OT account of Tzotzil AF verbs.

[9] Abbreviations in Tzotzil examples are: CP = completive aspect; ICP = incompletive aspect; E3 = ergative 3rd person; ENC = enclitic

'the man who's selling pineapple'

The restrictions on the AF form can be understood in terms of the interpretations available to the PLAIN, TRANSITIVE VERB (the unmarked form). The plain form is always preferred over the AF form, as long as it can be associated with the target interpretation. As in Takelma, the interpretations available to the plain form are determined by agreement and bias. Only interpretations unavailable to the plain form are expressed by the AF form.

Like *-k^hwa* in Takelma, the AF verb occurs only when both S and O are 3rd person because it is only in this case that verb agreement *underdetermines* the linkage of nominal arguments with grammatical functions. When both S and O are 3rd person, the AF verb occurs only in clauses that go against the bias for human, topical subjects. Again, the relation between the plain form and the AF form is a partial blocking relation: both forms occur, but the marked form is relegated to the expression of more marked interpretations.

In virtue of this relation, the AF verb serves a disambiguating function in some clauses. The examples in (17) differ only in that (a) has the plain transitive verb, while (b) has the AF verb.

(17) a. K'usi i-s-sibtas li antzetike?
what CP-E3-frighten the women
'What did the women frighten?'

b. K'usi i-sibtas-<u>on</u> li antzetike?
what CP-frighten-AF the women
'What frightened the women?'

A priori, (17a) – with a plain transitive verb – might be expected to be ambiguous between (*m1*) *what did the woman frighten?* and (*m2*) *what frightened the woman?* However, it has only the interpretation shown (*m1*). The reason is that its interpretation is narrowed by the Bias constraints, as shown in the evaluation in (18). Here the four candidates are formed by pairing the two forms in 17a,b with the two interpretations in 17a,b.

(18)

	*AF	SUB HUM	SUB TOP	*AF	SUB HUM	SUB TOP
☞f1 k'usi s-sibtas li antzetike *what E3-frighten the women*			*			
☞f2 k'usi sibtas-on li antzetike *what frighten-AF the women*				*	*	*
	☞ m1: what did the women frighten?			☞ m2: What frightened the women?		

Since *f1* (=17a) contains one human and one non-human argument, SUBJECT HUMAN prefers *m1* to *m2* as the optimal interpretation of *f1*. *f1* is also a better expression for *m1* than *f2* is, since *f2* violates *AF, a markedness constraint which penalizes the AF verb. <*f1, m1*> is thus Weak OT optimal (*), and blocks both <*f1, m2*> and <*f2, m1*> (the shaded quadrants). With <*f1, m2*> and <*f2, m1*> removed from competition, <*f2, m2*> emerges as a new Weak OT optimal pair (⌠): the marked form is paired with the marked interpretation. Even though <*f2, m2*> has the worst constraint profile among the four candidates, *f2* is the optimal expression for *m2*, and *m2* is the optimal interpretation of *f2*.[10]

While (17a, b) are each unambiguous in Tzotzil, it is not the case that all ambiguity in contexts of extraction from transitive clauses is resolved by the grammar of Tzotzil. The example in (19), for example, is ambiguous.

(19) Buch'u i-s-mil li Xune?
who CP-E3-kill the Juan
Who killed Juan? *and* Who did Juan kill?

In order for the Bias constraints to favor one interpretation over another, the two arguments must be differentiated by the Bias constraints – either with respect to animacy or topicality. In (19), both arguments are human, and neither is (by assumption) topical. Hence the constraints discussed up to this point will not rank the two interpretations. There is no *general* proscription against ambiguity at work in (17a,b). The ranking of candidate interpretations is determined by language-particular grammars.

7. Recoverability in head-marking grammar

A common feature of Takelma -*k^{h}wa* and Tzotzil Agent Focus is that both constructions are limited to clauses with two 3rd person arguments. In both cases, this restriction follows from the fact that in any other case, agreement fully determines the grammatical functions of nominal arguments so the extra marking provided by -*k^{h}wa* and the AF suffix is not needed.

The interpretive problem faced by clauses with two 3rd person arguments is of course not restricted to these languages. In general, head marking languages without much surface configurationality face the problem of recoverability in 3-3 clauses, but not in other clause types. It is fairly common in languages of the head-marking, low-configurationality type to find constructions or restrictions which are limited to 3-3 clauses (Nichols 1986).

Another restriction which tends to occur in the same languages that show animacy and topicality effects is a bias towards pronominal subjects and lexical objects. In a transitive clause with one overt nominal and one pronoun, this bias

[10] The restriction of AF forms to subject extraction contexts may be due to a higher-ranking faithfulness condition which is left inexplicit here. If so, the pair <*f2, m1*> would violate this constraint, and would have a worse profile than <*f2, m2*>. This would not affect the outcome.

leads to parsing the pronoun as subject and the overt nominal as object.[11] Gerdts (1988) observed this in Halkomelem, dubbing it the ONE NOMINAL CONSTRAINT.[12]

(20) a. ni qʷáqʷʌtʌs kʷθʌ swʌƛyʔqeʔ.
AUX club-TR-E3 DET man
He clubbed the man/*the man clubbed him. (p. 58)

b. !! ni qʷʌlʌtʌs ʉʌ sʉéniʔ.
AUX bake-TR-E3 DET woman
!!He baked the woman./*The woman baked it. (p. 58)

What appears to be the same constraint has been described for a number of other head-marking languages, including Chamorro (Chung 1984) and Navajo (Platero 1982, among other references). The constraint seems clearly related to recoverability, as it too is restricted in these languages to clauses with two 3rd person arguments. As above, some more complex construction is needed to express the interpretation which is blocked, for example, in (20a).

In short, there are various biases which are active in these head-marking languages, all restricted to clauses with two 3rd person arguments. In earlier work of my own, I suggested that these phenomena could be analyzed in terms of ABSTRACT OBVIATION (Aissen 1997). This account was inspired by Algonquian morphosyntax, where there is an overt morphological category of OBVIATION which is only relevant to 3rd persons, and is only significant when there are multiple 3rd persons in a discourse segment. In these obviation systems, 3rd persons are ranked according to properties like animacy and topicality, and this ranking determines various aspects of morphosyntax. In Aissen (1999b), I analyzed the agent focus construction in Tzotzil in terms of abstract obviation; Culy (2000) analyzed Takelma *-kʰwa* in the same terms. These accounts do not appeal to interpretive preferences, and do not relate these phenomena directly to the problem of recoverability that faces clauses with two 3rd person arguments. What I am suggesting here is that appeal to abstract obviation is not needed if Weak bidirectional OT is adopted. The restriction of Bias constraints to 3-3 clauses can be achieved in a Weak bidirectional OT without stipulation or abstract obviation since marked clauses with 1st and 2nd person arguments can afford to be expressed in the most economical fashion without risk of misinterpretation.

From this perspective, morphological obviation is one solution to the interpretive problem faced by head-marking languages in 3-3 clauses. But there

---

[11] One may ask whether the interpretation of (7a) (Takelma) might not follow from this constraint, rather than ones referencing animacy and topicality, as proposed above. Culy (2000) argues that such a constraint is neither necessary nor sufficient to account for the Taklema facts, while constraints based on animacy and topicality are.
[12] Gerdts uses "!̲ to mark forms which are structurally grammatical but semantically anomalous" (p. 9); TR = transitive suffix.

are other solutions, including differential morphology like Takelma *-k^{h}wa* and the Tzotzil AF form. Morphological obviation systems like those of Algonquian, and Kutenai are grammaticized systems which overextend morphological marking beyond where it is needed, strictly speaking, for the sake of recoverability. The kinds of systems discussed here are not overextended – they appear to be used exactly where they are needed for recoverability and not beyond.

References

Aissen, Judith. 1997. On the syntax of obviation. *Language* 73:705-750.

Aissen, Judith. 1999a. Markedness and subject choice in optimality theory. *Natural Language and Linguistic Theory* 17:673-711.

Aissen, Judith. 1999b. Agent focus and inverse in Tzotzil. *Language* 75:451-485.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21:435-483.

Aronoff, Mark. 1976. *Word formation in generative grammar.* Cambridge: The MIT Press.

Beaver, David and Hanjung Lee. to appear. Input-output mismatches in OT. In R. Blutner & H. Zeevat (eds.) *Optimality theory and pragmatics*: Palgrave/Macmillan.

Blutner, Reinhard. 1999. Some aspects of optimality in natural language interpretation. In H. d. Hoop & H. d. Swart (eds.) *Papers on optimality theoretic semantics.* Utrecht: Utrecht Institute of Linguistics.

Chung, Sandra. 1984. Identifiability and null objects in Chamorro. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society,* 116-130. Berkeley: Berkeley Linguistics Society.

Comrie, Bernard. 1989. *Language universals and linguistic typology.* Chicago: University of Chicago Press.

Culy, Christopher. 2000. An incorporated topic marker in Takelma. In M. Butt & T. H. King (eds.) *Proceedings of the LFG00 Conference*: CSLI On-Line Publications. http://csli-publications.stanford.edu/LFG/5/lfg00culy.pdf.

DeLancey, Scott. 1981. An interpretation of split ergativity. *Language* 57:626-657.

Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland.* Cambridge: Cambridge University Press.

Gerdts, Donna. 1988. *Object and absolutive in Halkomelem Salish.* New York: Garland.

Golla, Victor ed. 1990. *The collected works of Edward Sapir.* vol. VIII. Berlin: Mouton de Gruyter.

Grimshaw, Jane. 1997. Projections, heads, and optimality. *Linguistic Inquiry* 28:373-422.

Hendriks, Petra and Helen de Hoop. 2001. Optimality theoretic semantics. *Linguistics and Philosophy* 24:1-32.

Horn, Laurence. 1989. *A natural history of negation.* Chicago: University of Chicago Press.

Jelinek, Eloise and Richard Demers. 1983. The agent hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics* 49:167-185.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In Linguistic Soc. of Korea (ed.) *Linguistics in the morning calm.* Seoul: Hanshin Publishing Co.

McCawley, James. 1978. Conversational implicature and the lexicon. In P. Cole (ed.) *Syntax and Semantics 9, Pragmatics,* 245-259. New York: Academic Press.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62:56-119.

Platero, Paul. 1982. Missing noun phrases and grammatical relations in Navajo. *International Journal of American Linguistics* 48:286-305.

Sapir, Edward. 1922. The Takelma language of southwestern Oregon. In F. Boas (ed.) *Bulletin of American Ethnology: Handbook of American Indian Languages,* 1-296. Washington, D.C.: Government Printing Office.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.) *Grammatical categories in Australian languages,* 112-171. Canberra: Australian Institute of Aboriginal Studies.

Stiebels, Barbara. 2003. Agent focus in Mayan languages. Ms. Berlin.

Wilson, Colin. 2001. Bidirectional optimization and the theory of anaphora. In G. Legendre, J. Grimshaw & S. Vikner (eds.) *Optimality-theoretic syntax,* 465-508. Cambridge, MA: The MIT Press.

Judith Aissen
Linguistics Department
UCSC
Santa Cruz, CA 95064

aissen@ucsc.edu

Does *most* mean 'more than half'?

MIRA ARIEL
Tel Aviv University

1. *Most* and *more than half*

Linguists have analyzed the lexical meaning of *most* as 'more than half', assuming that the semantics of *most* is only lower-bounded, just like that of *more than half*: Both expressions cannot denote anything less than 50% plus something (51% for short). Crucially, this means that they are not upper-bounded, and can therefore denote 'all'. Researchers do not, however, assume that *most* actually gets interpreted as 'all' in the default case. Rather, they assume that an upper bound is imposed extralinguistically. The Neo-Griceans in particular have proposed that a scalar Generalized Conversational implicature 'not all' is generated by default, creating a conveyed meaning of 'more than half but not all' (see Horn 1972, 1989, Levinson 2000).[1] In Ariel 2004 I argue at length against such a view. I cite examples from natural conversations in English, where the generation of a 'not all' implicature seems highly unmotivated, since it would defeat the speaker's purpose in using a *most* utterance. I further argue that the cases where the implicature is not generated constitute the majority of the uses of *most*.

If we maintain the received lower-bound-only semantic analysis, and at the same time accept my conclusion that a default 'not all' implicature is not usually generated, neither the semantics nor the pragmatics of *most* can account for the upper bound placed on it. Fred Landman (2000: Chapter 7, p.c.), however, proposes that an upper bound is in effect imposed in the default case because statistically, the chances are simply very slim that precisely a 100% value will be selected by addressees.[2] Such an analysis can then preserve the received semantic meaning view, without having to rely on a default scalar implicature, which I claim is not in fact generated in discourse. Under such a view we expect no difference between *most* and *more than half*. Both mean 'more than half', and should equally (though rarely) denote 100%.

---

[1] Most claims in the literature pertain to *some*, rather than to *most*, but the assumption is that *most* too is only lower bounded semantically, and upper bounded pragmatically.

[2] I thank Fred Landman for raising pertinent challenges, to which this paper is in part a response.

I would like to argue that *most* and *more than half* are semantically distinct. Only the latter means 'more than half'. It is its lexical meaning, and not merely a low probability, that blocks the 100% interpretation for *most*. In order to support a distinction between *most* and *more than half*, I examined speakers' interpretations of these two expressions in a questionnaire. I presented subjects with two types of questions. The first included sentences containing Hebrew *rov* 'most' or *yoter mexaci/lemaala mimaxacit* 'more than half', for the most part followed by multiple choice questions as to potential values which can be assigned to *most* or to *more than half*.[3] I started with a pilot study, followed by the main study, and the results below combine the results from both studies whenever the questions remained unchanged. One group (32 students of the Hebrew Department of Tel Aviv University, all native speakers of Hebrew) were asked about *most*, and another (19 different students of the Hebrew Department at Tel Aviv University, all native speakers of Hebrew) were asked about *more than half*.[4] To tap their array of intuitions, participants were explicitly asked to select all possible answers, including answers that were highly unlikely (e.g., (3)b, c, (4)a, below for *most*). In a second, smaller set of questions, I asked subjects to determine the interpretation of a discourse anaphoric pronoun. The antecedent was *most*, *more than half*, or some number (in order to compare between *most* and numbers). Subjects were then asked to select only one answer (e.g., (8), (9) below).

We should bear in mind, however, that native speakers' intuitions do not distinguish semantic from pragmatic meanings. Rather, speakers' judgments can only attest to conveyed meanings directly. Hence, not every difference we find between the interpretation of *most* and *more than half* can automatically be attributed to their semantics. The difference may only be pragmatic, since identical semantic meanings can give rise to different pragmatic interpretations, due to the Maxim of Manner. In order to distinguish between pragmatic and semantic differences I have created situations in which subjects were encouraged to select as the conveyed meaning of *most* and *more than half* meanings which are pragmatically dispreferred, although semantically possible.

All in all, I will argue that *most* and *more than half* share a lower semantic bound (51%) and an extralinguistic tendency of avoiding extremely high values. They differ pragmatically in that *more than half* tends to receive lower majority interpretations, whereas *most* tends to receive larger majority interpretations. Most importantly, they differ regarding the upper bound. *More than half* is not, but *most* is, semantically upper bounded.

---

[3] Based on a large corpus of conversational English data and many Hebrew examples I collected I feel quite confident that there is no difference between the Hebrew examples tested for and their English counterparts (where *most* is the appropriate translation). But of course, such questionnaires should be administered to speakers of English and other languages as well.

[4] None of the students has taken a Linguistics Department course in semantics (where they would have been taught the received view that *most*, just like *more than half*, is only lower bounded lexically).

2. Pragmatic differences between *most* and *more than half*

Consider the following:

(1) **Most** of the ladies and **more than half** of the gentlemen wore evening clothes (Sinclair Lewis, *It can't happen here*, McCawley's example 14.1.5, p. 427).

As McCawley 1981: 427 explains, this quote "strongly suggests that a greater proportion of ladies than of gentlemen were dressed in evening clothes". It seems that *more than half*, just because it takes the half point as a reference point, is normally interpreted as above, but still close to 50%. This is why it is taken as a lower majority than *most*. *Most*, on the other hand, tends to denote a significant (though not an overwhelming) majority, which is quite natural, given its etymology. A noncategorical skewing in the acceptance rates of higher values for *most* and lower values *more than half* should then be seen as evidence for a pragmatic distinction between the two expressions.

To test for this pragmatic tendency in my questionnaire, subjects had to select **all** values that the speaker could have in mind when using *most/more than half.* (2) and (3) are two such test examples:[5]

(2) He lived **most/more than half** of his life in Israel.
Which of the cases could the speaker have in mind?
a. 60% of his life b. 75% of his life c. 48% of his life
d. 90% of his life e. none of the above.

(3) **Most/more than half** of the cars in this shop are defective.
How many cars could the speaker mean?
a. 31% of the cars b. 51% of the cars c. 99% of the cars
d. 85% of the cars e. none of the above.

The results obtained in the study confirm the pragmatic difference between *most* and *more than half* noted by McCawley. Table (1) presents the acceptance rates for the lower values: 51%, 60% and 75%:

Table 1. Acceptance of 51-75% values for *most* and *more than half*

Value	***Most***	***More than half***
51%	47/64=73.4%	35/38=92.1%
60%	21/32=65.6%	19/19=100%
75%	30/40=75%	25/26=96.15%
51+60%+75%	98/136=72.1%	79/83=95.2%

[5] Note that each group only saw one of the quantifiers. Also, boldface was not used in the questionnaire.

Whereas there is no significant difference between *most* and *more than half* for the total range of 51-99%, there is for 51-75%:[6] There were over 30% more confirmations for these values for *more than half* than for *most.*[7]

When we look at 80% and 85%, the acceptance rate for *more than half* remains very high. But the acceptance rate for *most* rises, so there is no difference between the two expressions (there are only 1.04 more positive responses (proportionately) for *more than half* than for *most*):[8]

Table 2. Acceptance of 80%, 85% values for *most* and *more than half*

Value	***Most***	***More than half***
80%	29/32=90.6%	18/19=94.7%
85%	23/24=95.8%	12/12=100%
80%+85%	52/56=92.9%	30/31=96.8%

Once we examine rates of acceptance for very high values, it is *most* which receives a higher acceptance rate. Overall, there are (proportionately) 1.39 more confirmations for *most* than for *more than half* in this range:

Table 3. Acceptance of 90, 99% values for *most* and *more than half*

Value	***Most***	***More than half***
90%	27/32=84.4%	10/19=52.6%
99%	48/64=75%	22/38=57.9%
90%+ 99%	75/96=78.1%	32/57=56.1%

Note, however, that acceptance rates decrease for both expressions for these extremely high percentages, where *virtually all* would have been appropriate. This is especially true for *more than half* and for 99% for *most.* The differences between *most* and *more than half* show up at small and at very large majorities: Subjects consistently preferred to interpret higher percentages for *most* and lower percentages for *more than half.* It seems that *more than half* has a wider range of high acceptance than *most*: It receives extremely high rates of acceptance for 51%-85% (109/114, 95.6%). It then drops rather sharply for the very high values, 90%, 99% (56.1%). *Most,* on the other hand, peaks mostly at 80%, 85% (92.9%). It is almost equally low for the low range and for the extremely high range of 99% (72.1% vs. 75% respectively).

All in all, the questionnaire results confirm that *more than half* is associated with relatively lower majorities than *most* is. Since none of the differences between *most* and *more than half* is close to absolute, it is reasonable to attribute the differences to different pragmatic associations attached to the two expressions:

[6] I thank Gila Batori for the statistical analyses in this paper.

[7] Note that values were repeated in 1-3 questions, and were therefore responded to by different numbers of subjects. Hence the use of proportionate numbers everywhere.

[8] Due to the small number of different questions posed for the same value, it is impossible to perform statistical analyses of variance on many of the values.

More than half emphasizes the relevance of the half point, whereas *most* emphasizes the significance of the majority involved. In section 3 I show that another difference between *most* and *more than half* is categorical, and therefore constitutes a semantic difference.

3. Determining the semantic boundaries on *most* and *more than half*

Let us now examine the semantic boundaries on *most* and *more than half,* and see whether they are identical (namely, lower bounded only). Specifically, regarding the upper bound, I will argue that it is pragmatic for *more than half,* but semantic for *most*.

We first ascertain the semantic lower bound on *most* and *more than half.* Expectations here were that subjects would absolutely avoid any value lower than 51% for either expression. Findings demonstrate that the lower bound on *most* and *more than half* is adhered to, but not quite as absolutely. A very low percentage of responses confirmed values lower than 51% (0%, 15%, 20%, 28%, 30%, 31%, 40%, 46%, 48% or 49%, 50%, all potential answers) for both expressions. Surprisingly, however, the ban against interpreting *most* as lower bounded is weaker than that on *more than half*: 6% of the responses below 51% were accepted for *most* (22/368), as opposed to half that rate for *more than half* (6/214, 2.8%, a statistically significant difference). Nonetheless, an overwhelming majority of the responses rejected values lower than 51% for both *most* (94%) and *more than half* (97.2%).

To see that the lower bound is indeed an absolute boundary (in as much as there are absolute boundaries on linguistic meanings in natural language) we can compare the acceptance rates for the minimally different 49%, 50% and 51%:

Table 4. Acceptance of values 49%, 50%, 51% for *most* and *more than half*

Value	***Most***	***More than half***
49%	3/32=9.4%	1/19=5.3%
50%	7/64=10.9%	3/38=7.9%
51%	47/64=73.4%	35/38=92.1%

There is an increase in acceptance for 50% over 49%, but it is relatively mild (1.17 times more for *most*, 1.5 times more for *more than half*). The increase for 51% over 50%, on the other hand, is dramatic: 6.7 times more for *most*, 11.7 times more for *more than half.* No such gaps were observed between the higher values discussed in section 2, which were more than minimally different from each other (e.g., between 85% and 80%, where the gap in favor of the higher rate is 1.06 for both *most* and *more than half*). In sum, while statistically, the ban against values lower than 51% is not total, I suggest that it counts as linguistically categorical.

We now move on to the upper bound. According to Landman, subjects should be equally reluctant to select 100% for *most* and for *more than half.* In the light of the pragmatic tendency to associate *most* with higher values, higher rates of 100%

acceptance are expected for *most* than for *more than half*. My expectations, however, were that although subjects would be quite reluctant to accept 100% as an answer for *more than half* due to the pragmatic tendency observed above, some subjects might select it for some questions, because it is after all a legitimate semantic interpretation. For *most*, I expected that subjects would categorically reject 100% , i.e., at the same (almost) absolute rate that they rejected 50%. Recall that subjects were encouraged to circle as many responses as they could, even if there was only a remote chance that the speaker thought they were possible. In order to encourage subjects to select extensions which are claimed to be semantically possible, despite a dispreferred pragmatic status (e.g., 51%, 99% and possibly 100% for *most*), I presented subjects with sentences such as (4) (with 100%), where the other candidate answers are semantically ruled out:

(4) **Most/more than half** of the students in the class were born in 1970.
How many students could the speaker mean?
a. 100% of the students b. 20% of the students c. 50% of the students
d. 49% of the students e. None of the above.

I made sure that 51%, 99% and 100% all occurred as answers in two types of cases: Where the other candidates were plausible (e.g., (3)), or semantically impossible (as in (4)). The idea was that although subjects were instructed to choose a maximal number of answers, regardless of how plausible they were, they might still be reluctant to choose a dispreferred option when a more plausible one was available. They might, however, be driven to select the dispreferred option once their other choices were either semantically inappropriate or 'none of the above'. This indeed was the case for 51% for *most*, for example: 21/32 (65.6%) accepted 51% when other semantic options were available, 26/32 (81.25%) accepted 51% when no other semantically appropriate choice was available.

A similar attempt to induce subjects to accept a 100% value for *most* was not as successful, however. It seems that despite the fact that *most* is associated with higher values, it simply cannot denote 100%, and despite the fact that more *than half* is pragmatically associated with lower values than *most*, it can denote the highest value, 100%. Even when I tried to encourage subjects to select 100% as a possible denotation for *most/more than half* (as in (4)), where, moreover, students in the same grade could conceivably have all been born in the same year, only a handful of the subjects tested on *most* chose the 100% option (3/32, 9.4%). In fact, there was an equal number of 49% responses here, and even more 50% responses (5/32, 15.6%). At least some subjects view the upper bound (99%) as more rigid than the lower bound (51%). Most subjects, however, opted for the "none of the above" answer (26/32, 81.25%), equally avoiding the violation of either boundary.

Now, if this is also the case for *more than half*, we could perhaps try to explain the findings by reference to some very strong pragmatic tendency. But this is not the case. Whereas 11/19 (57.9%) students tested on the *more than half* parallel example also chose "none of the above", 8/19 (42.1%) did select 100% as a possibility here (4.5 times more than *most*). When in another test sentence (*Most/more than half of high school students drink alcohol*) the 100% answer was presented alongside a more pragmatically suitable candidate answer (80%), only 2/32 (6.25%) accepted the 100% answer for *most* (with an equal number of responses confirming 50% here), but still, 7/19 (36.8%) agreed that *more than half* could denote 100% (5.9 times more accepting 100% for *more than half*).

Table (5) presents the acceptance rates for 99%, 100% and 50% for *most* and for *more than half* in the above two questions.

Table 5. Rates of acceptance for 99%, 100% for *most* and *more than half*

Value	***Most***	***More than half***
99%	51/64=79.7%	22/38=57.9%
100%	5/64=7.8%	15/38=39.5%
50%	7/64=10.9%	3/38=7.9%

While the acceptance rate for 100% is significantly lower than for 99% for *more than half* (a 1.5 times reduction), the difference for *most* is far more dramatic. There are more than 10 responses confirming 99% per one confirming 100% here. In fact, this gap is even larger than the gap (of 6.7) between 51% and 50% for *most*, which we have taken above as linguistically categorical. Note also that acceptance rates for 100% are actually 1.4 times lower than for 50% for *most*. In contrast, for *more than half* 100% is accepted 5 times more than 50%.

Finally, in another attempt to encourage subjects to allow for very high values for both expressions, I asked about *rov acum/harbe lemaala mixeci* 'an **overwhelming** majority'/'**a lot** more than half' (*of the students passed the test*). In addition, as options I only listed extremely high percentages: 97%, 98%, 99% and 100% ('none of the above' was not an option here), thus forcing subjects to suppress their pragmatic preferences to avoid extremely high percentages for both expressions.[9] Table (6) presents the results:

Table 6. High percentages as options for *most* and *more than half*

	100%	**99%**	**98%**	**97%**
Most	1/32=3.1%	29/32=90.6%	29/32=90.6%	31/32=96.9%
More than half	6/18=33.3%	15/18=83.3%	15/18=83.3%	18/18=100%

First, note that indeed, subjects were now willing to select the very high values of 97%-99% for both expressions. They still manifest a mild decrease in their acceptance of higher percentages for both *most* and *more than half* (as we saw above, both expressions showed a decrease for extremely high values). Thus, 98% and 99% are selected less often than 97% for both expressions. Now, the decrease for 100% over 99% is quite significant for *more than half* (2.5 times less for

[9] I thank John Du Bois for suggesting to me this set of answers.

100% than for 99%), but this gap is by far smaller than the one for *most*: 29 times less confirmations for 100% for *most*. Whereas there are 1.09 more confirmations for 99% for *most* than for *more than half* in this question, results are switched around, and dramatically so for 100%. Here there are 11.6 times more confirmations for *more than half* than for *most*.

All in all, subjects did prefer to view both *most* and *more than half* as upper bounded. However, this tendency is significantly stronger for *most*. A comparison between the positive answers for all three 99% vs. 100% questions is instructive:

Table 7. 99%, 100% in all 3 questions

	100%	**99%**
Most	6/96=6.25%	80/96=83.3%
More than half	21/56=37.5%	37/56=66.1%

In conformity with the pragmatic tendency noted in section 2, the percentage of answers confirming 99% for *most* is higher than that for *more than half* (1.25 times more). But preferences are reversed for 100%. For *most* the drop is virtually absolute: There are 13.3 times more confirmations for 99% than for 100%. The drop is much milder for *more than half* (by 1.75). While subjects avoided the association of *more than half* with 100% in close to two thirds of all the relevant questions (62.5%), more than a third of the answers (37.5%) shows it as a legitimate interpretation of *more than half* (and *a lot more than half*) (6 times more than for *most* and *an overwhelming majority*).

Now, can this clear statistical difference between *most* and *more than half* re the upper bound be accounted for pragmatically? I think not. Assuming, as we should, that *more than half* is oriented towards the half point, whereas *most* is pragmatically associated with a more significant majority, we cannot explain the results in section 3. Since in general, it is *most* that is pragmatically taken to denote a larger proportion than *more than half*, a tendency maintained up to the 99% level, we cannot apply the same pragmatic analysis to also explain the opposite finding, namely, the remarkably lower acceptance of the highest percentage, 100%, for *most*. According to the pragmatic difference noted in section 2, precisely the opposite would have been expected: Higher proportions should be expected for the quantifier usually reserved for larger majorities, namely *most*, rather than *more than half*. The conclusion must be that for *most*, 100% is categorically different from 99%.

In conclusion, the ban on viewing *most* as denoting 100% is absolute (93.75%). Recall that the ban against 50% satisfying *most* is slightly lower actually – 89.1%. I therefore suggest that the upper bound constraint is semantic for *most*. Refusing to view *more than half* as denoting 100% is only a strong pragmatic preference, applied in almost two thirds of the cases. I claim that the only way to account for the different rates of avoiding 100% options is by positing a semantic difference between *most* and *more than half*. Whereas *more than half* is not, *most* is, lexically upper bounded. We now turn to different questionnaire results, which corroborate the same conclusion re the semantics of *most*. We here compare *most* with numbers.

4. An argument from discourse anaphora

Section 4 attempts to demonstrate that a test which has been used to decide between unilateral (lower bounded only) and bilateral (also upper bounded) meanings for numbers can also be adduced in support of a bilateral meaning for *most*. Kadmon (1987, 2001) proposed a discourse anaphora test in deciding whether the numbers are lexically only unilateral or bilateral. She provides the following example (Kadmon 2001: 72, ex. 14):

(5) **Eleven** kids walked into the room. **They** were making an awful lot of noise.

Kadmon argues that whereas *eleven* above is compatible with there being more than 11 kids who walked into the room (truth-conditionally, 'eleven' is equivalent to 'at least eleven'), *they* cannot refer to anything but 'exactly eleven'. All things being equal, the only unique antecedent that the pronoun can pick is the conveyed meaning 'exactly eleven kids', created by adding on a scalar implicature ('no more than 11 kids') to the lexical meaning of *eleven kids* 'a set of eleven kids'. If it is known that, say, 12 kids walked into the room, the discourse in (5) is infelicitous, because there is no unique referent that *they* can refer to (the scalar implicature cannot be generated in this case).[10] This is not the case for *at least eleven kids....*, which does constitute a unique antecedent for a following discourse anaphoric pronoun. Thus, *they* (6) refers to all the kids that walked into the room, regardless of whether there were 11 or 12 of them:

(6) **At least eleven** kids walked into the room. **They** were making an awful lot of noise.

Based on these observations, Nirit Kadmon (p.c.) and Fred Landman (p.c.) propose that if *most* is only lower-bounded, it should pattern with *at least n*, rather than with *n*. Thus, anaphora resolution for *most* should resemble the one in (6), and not in (5). For example, for (7), Landman predicts that *them* will refer to all the defective Hondas, should it be the case that all Hondas are defective:

(7) **Most** Hondas were defective. They took **them** out of the shop.

If so, examples (5) and (7) testify to a difference re upper bound between the numbers and the quantifier *most*. Since *they*, when coreferent with a number seems to refer to that number exactly, the semantic meaning of the numbers must be bilateral. Since *they*, when coreferent with *most*, seems to pick only a lower bounded referent, the semantic meaning of *most* must be unilateral.

[10] Unless a unique referent can pragmatically be created in some special way.

I would like to question some of these intuitions. The findings below testify that discourse anaphoric pronouns are not as restricted as Kadmon and Landman envision. Pronouns seem to rather easily select various pragmatically enriched interpretations as antecedents.[11] In the same questionnaire referred to above, I also asked the subjects to choose a referent for such anaphoric expressions. Here is an example:

(8) Ruti told me that **most** of the teachers are interested in changing the school principal in Karmiel. "**They** even signed a petition against him, which was sent to the Minister of Education", she added.

Question: It became apparent that **all** the school teachers are interested in changing the principal. Who are those that Ruti meant that "**They** even signed a petition against him, which was sent to the Minister of Education"?

Answers: A. Between 51% and 99% of the school teachers
Or: B. 100% of the school teachers Or: C. Impossible to know.

One subject (out of 24, 4.2%) said that *They* refers to 'all' (answer B), in line with Landman's prediction. 7/24 (29.2%) chose Answer C, which is what Kadmon predicts for the numbers. Crucially, two thirds of the subjects (16/24) chose 'most but not all' as the intended referent (Answer A). This is clearly contra Landman's claim. If *most* can denote 'all' and we know that 'all' is the case, the pronoun should have referred to 'all'. But it didn't in most cases.

Let us now make sure that the above results are similar enough to parallel results re the numbers. Indeed, the *most* results are quite similar to the results I obtained for the number *14* (asked about in two additional sets of questions, one per each group of subjects). Since the results from the two questions here are quite similar, I will combine their results. Here is one version of the question:

(9) Before she left to go home the substitute teacher reported to the school principal an insignificant disciplinary problem that she had. **14** of the 9th grade students left the classroom late for recess. **They** misbehaved and made a lot of noise, but she decided not to punish them.

Question: It was found out that **15** of the students left the classroom late. How many students, in your opinion, did the substitute teacher refer to when she said "They misbehaved and made a lot of noise"?

Answers: A. 15 Or: B. 14 Or: C. Impossible to know.

In conformity with Kadmon's predictions, 26/50 (52%) chose 'impossible to know' whether *they* refers to '14' or to '15' when the antecedent was *14* (Cf. 29.2% for *most* in Question (8)). However, 21/50 (42%) said that *they* refers to '14' (Cf. 66.7% 'most but not all' for *most*). Finally, only 3/50 (6%) said that *they*

---

[11] Kadmon concedes that there are pragmatic enrichments which change semantic meanings (accommodations), but finds them less than fully acceptable.

can refer to '15 students' (Cf. 4.2% 'all' for *most*). Interestingly, one subject (not counted above) chose '14' as the appropriate answer "in principle" and 'impossible to know' as the appropriate answer "in an actual situation".

The answers above demonstrate that subjects vary in their interpretations of a pronoun whose antecedent is either a number or the quantifier *most*. In both cases (especially for *most*) there was a clear preference for the (also) upper bounded value over the lower bounded meaning, even in the absence of a scalar implicature, but this was not absolutely so. Now, in the questions quoted so far (in (8) and (9)), the context I created encouraged the subjects to stick to the upper bound for the anaphoric expression, because the activity described in the sentence containing the anaphoric *they* could very well be true for only a subset of the entities for which the activity described in the antecedent sentence was true. (While all the teachers may be interested in changing the principal, not necessarily all signed the petition; while 15 students left class late, not necessarily all made a lot of noise.) In order to show that it is indeed context that determines the interpretation of discourse anaphors, I constructed another set of questions, where the context was biased towards interpreting the antecedent as only lower bounded, and towards viewing the two sets as identical. Here I expected the anaphoric pronoun to pick its referent from what I consider the pragmatically enriched 'at least' interpretation of the antecedent (*most* or number).

Indeed, this is precisely what happened for both the number (*61*) and for *most*. In the following, the 'at least 61/most' reading is strongly preferred as the contextually appropriate interpretation for *61/most*, and in addition, it is extremely likely that an identical set is intended in both sentences. Accordingly, subjects tended to choose an 'at least' reading for the anaphoric *they* here. Consider first the *most* question:[12]

(10) It is not necessary for all 120 Knesset members to be present for the Knesset to convene. The law requires that **most** Knesset members participate in the assembly discussion tomorrow. **They** will be asked to vote on a series of social laws.

Question: Who are the "they" who "will be asked to vote on a series of social laws"?

Answers: A. All the Knesset members who will participate in the assembly discussion tomorrow, that is, no less than 61 and up to 120 Knesset members, including 120 Knesset members.

Or: B. All the Knesset members who will participate in the assembly discussion tomorrow, that is, no less than 61 and no more than 119 Knesset members.

A minority (5/30, 16.7%) still insisted that the anaphoric *they* here only refers to an upper-bound 'most but not all' (answer B). Since this answer is pragmatically

[12] Knesset is the name of the Israeli parliament.

inappropriate in this case, it must be the semantics of *most* which accounts for this choice. But the great majority of the respondents chose answer A (25/30, 83.3%), assigning *they* an 'at least most' reading. Similar results were obtained when I asked about *61* (I substituted *61* for *most* in the above). In addition, Answer B was changed to:

(11) **Answer B**: All the Knesset members who will participate in the assembly discussion tomorrow, that is, no less and no more than 61 Knesset members.

In this case, an even lower proportion of subjects (2/19, 10.5%) chose an upper-bound 'exactly 61' (Answer B), which is pragmatically inappropriate. 17/19 (89.5%) chose what I consider to be the pragmatically enriched 'at least' interpretation for the pronoun (Answer A). A comparison between *most* and *61* as antecedents for *they* in (10) and its counterpart shows that they pattern quite similarly, the 'at least' reading being the preferred one in this case. The conclusion must be that the anaphora test does not automatically testify to the lexical meaning of the numbers and *most*. Rather, it is the pragmatically enriched meaning (of *most* and *61*) which tends to serve as antecedent for the following anaphoric pronoun. We thus have one case where 'all'/'a higher number' was categorically rejected ((8) and (9)), and one case where it was selected in the great majority of the cases (10). In fact, in another context I created for a *most* question (not here quoted), I got an intermediate result, where an 'all' antecedent was selected in 31.8% of the responses (and 'most but not all' in 68.2%).

As further support for the context dependence of anaphoric resolutions, consider the results for *more than half* as a discourse antecedent for *they*. In the *more than half* counterpart of (8), no subject chose 100%, all subjects preferring '51-99%'. Respondents then avoided 'all', despite the semantic compatibility of 'all' with *more than half*, and despite the reality of the situation, where 'all' was presented as true for the antecedent proposition. So, even the fact that an expression is only lower bounded, as *more than half* is, does not guarantee that the anaphoric pronoun refer to 'all'.

Now, if context determines the interpretation of the pronoun, how can we know what the lexical contribution of the antecedent is, as opposed to pragmatic adaptation? I believe that the context which biases us towards possibly distinguishing the two sets, is the crucial one. Under such circumstances, the pronoun picks out the **lexical** meaning of the antecedent as the pragmatically plausible interpretation for the pronoun, despite the explicit claim that the antecedent refers to 'all' in reality. The reason is that the lexical meaning is all the addressee can be confident that the speaker intended for sure. Under such circumstances we saw that the majority of the subjects chose an upper bounded interpretation. Crucially, subjects made the same choice of upper bounded interpretation for *most* and for the numbers, testifying that the lexical meaning of *most* must be upper bounded, just like the lexical meaning of the numbers is (see Carston 1990, 1998, Geurts 1998 and references therein).[13]

In conclusion, the percentage estimates and the discourse anaphora questions point that *most* is lexically upper bounded. Recall that subjects virtually always refused to allow that *most* denoted 'all', even when no other legitimate value was available to them. This fact cannot stem from some absolute aversion to selecting 100%, since subjects did select 100% for *more than half* in a substantial minority of the cases. I have also argued that the avoidance of 100% for *most* but not for *more than half* cannot be explained by the pragmatic difference between the two expressions, since for values lower than 100% (up to 99%), it is *most* that selects the higher percentages, *more than half* showing a clear preference for smaller majorities. The discourse anaphora test, when applied to cases where there is a potential difference between the antecedent set and the anaphoric set, supports the percentage evaluation test in that it too demonstrates that the lexical meaning of *most* must be upper bounded. Only by assuming such an upper bound for *most* can we explain the fact that very often subjects took the anaphoric pronoun to refer to an upper bounded quantity, even though they were told that 'all' was the case for the antecedent.

Such choices point to the semantic, rather than extralinguistic status of the upper bound on *most*. Each of the three extralinguistic solutions fails to account for the ban on 100% for *most*. The fact that *most* received an upper bound in the absence of a scalar implicature (in the discourse anaphora questions) demonstrates that the upper bound is independent of such implicatures (and see Ariel 2004), the fact that *more than half* and *most* sharply diverge from their general pragmatic patterns when 100% is concerned shows that pragmatic tendencies (due to the Maxim of Manner) cannot here provide the explanation, and finally, the fact that the gap in acceptance of 100% and 99% was so dramatic for *most* demonstrates that pure probabilities cannot account for it either. The conclusion must be that *most* cannot mean 'more than half', because its semantics incorporates both lower and upper bounds.[14]

References

Ariel, Mira. 2004. Most. *Language 80.*

Carston, Robyn. 1990. Quantity maxims and generalized implicature. *UCL Working Papers in Linguistics* 2: 1-31. Revised version (1995) in *Lingua* 96. 213-44.

[13] Note that this is not Kadmon's position. Her explanation for the findings where the pronoun was interpreted as an upper bounded number is that despite the claim re reality, a scalar implicature was attributed to the speaker.

[14] On the precise nature of the upper bound see Ariel 2004.

Carston, Robyn. 1998. Informativeness, relevance and scalar implicature. In Robyn Carston & Seiji Uchida (eds.) *Relevance theory: Applications and interpretations,* 179-236. Amsterdam: Benjamins.

Geurts, Bart. 1998. Scalars. In P. Ludewig & B. Geurts (eds.) *Lexikalische Semantik aus kognitiver Sicht*, 95-117. Tübingen: Gunter Narr Verlag.

Horn, Laurence R. 1972. *On the semantic properties of logical operators in English.* Mimeo, Indiana University Linguistics Club, Bloomington, IN.

Horn, Laurence R. 1989. *A natural history of negation.* Chicago: University of Chicago Press.

Kadmon, Nirit. 1987. *On unique and non-unique reference and asymmetric quantification.* Ph.D. dissertation, University of Massachusetts, Amherst. Published 1992. Garland Press.

Kadmon, Nirit. 2001. *Formal pragmatics*. Oxford: Blackwell.

Landman, Fred. 2000. *Events and plurality*. Dordrecht: Kluwer.

Levinson, Stephen C. 2000. *Presumptive meanings*: *The theory of generalized conversational implicature*. Cambridge, MA: The MIT Press.

McCawley, James D. 1981. *Everything that linguists have always wanted to know about logic but were ashamed to ask.* Oxford: Blackwell.

Linguistics
Tel Aviv University
Tel Aviv, 69978
Israel

mariel@post.tau.ac.il

Functional Specialization in Tone Perception: Evidence from Dichotic Listening in Yalálag Zapotec[1]

HERIBERTO AVELINO
Department of Linguistics, UCLA

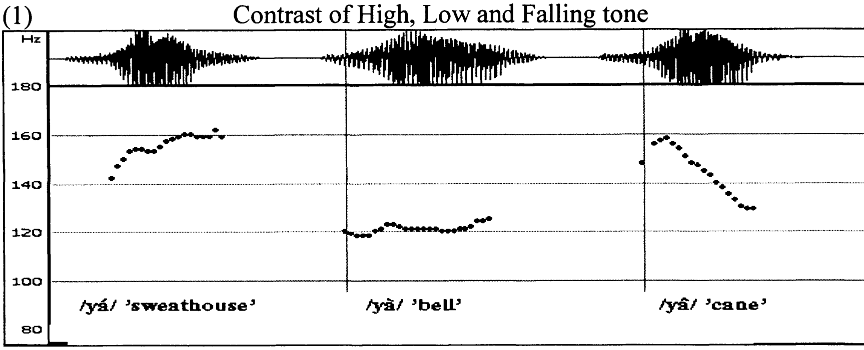
1 Introduction. An important tradition of clinical and experimental studies suggests that there is a left hemisphere specialization substrate for linguistic functions, while the right hemisphere is less involved in linguistic tasks (e.g. Fitch *et al.* 1997, Kimura 1961, Moffat and Hampson 2000, among others). This model of asymmetrical functions in the human brain, particularly the linguistic pre-potence of the left hemisphere, has been grounded mainly on clinical evidence. Findings of many studies point out that the likelihood of aphasia following damage of the left hemisphere of the brain is far higher than in cases where the lesion has affected the right hemisphere (e.g. Corina 1998, Penfield and Roberts 1959). More recent evidence from non-aphasic populations suggests a similar conclusion. In particular, dichotic listening studies in normal subjects and commissurotomized patients have demonstrated a consistent right ear advantage (REA) in linguistics stimuli, an observation that has been interpreted as an indication of left hemisphere superiority in processing such stimuli (Hugdahl 1988, Kimura 1961, 1967, Zaidel 1976, among others). Some studies have indicated that the right hemisphere is particularly involved in processing stimuli based on fundamental frequency, such as pure tones and intonation (Blumstein *et al.* 1974, Goodglass and Calderon 1977, Perkins *et al.* 1996). However, other studies of pitch perception in languages which use phonemic contrasts of tone (Gandour *et al.* 1988, Hugdahl *et al.* 1999) suggest that there is a left hemisphere dominance in the perception of phonemic tone. Experiments that have used a dichotic listening technique to examine the processing of linguistic vs. non-linguistic use of pitch are particularly relevant for the study described in this paper (Moen 1993, Van Lancker 1980, Van Lancker and Fromkin 1973, 1978, Wang *et al.* 2001) have reported a right ear advantage for discriminating tone. In agreement with the generally accepted model of hemispheric specialization for language, these results have been interpreted indicating that the perception of

[1] Thanks to Daria Allende, Ana Daisy Alonso, Estela Canseco, Mario Molina and specially to José Bollo my teacher of Zapotec, without their collaboration this work could have not been possible. I also want to thank Susan Curtiss, Christina Esposito, Sahyang Kim, Jody Kreiman, Peter Ladefoged and Pam Munro for the valuable comments and criticisms to earlier versions of this paper. All errors remain my own. The present research was supported by grants from UC-MEXUS/CONACYT which are gratefully acknowledged.

lexical tone is lateralized to the left hemisphere. Because the acoustic parameter fundamental frequency (f0) underlies a range of linguistic and non-linguistic functions, the study of functional specialization for the perception of f0 in languages with contrastive use of pitch presents crucial evidence to understand how tonal stimuli is processed in the brain. In this context, the present study would like to contribute towards the major research agenda on understanding of the neural correlates of language. This paper reports an investigation of the patterns of hemispheric specialization for tone perception in native listeners of Yalálag Zapotec (a tone language). Using a dichotic listening technique, three issues were investigated: the pattern of lateralization of pitch perception of phonemic tone in Yalálag Zapotec, in contrast to pitch perception of non-linguistic stimuli; the role of selective attention in the perception of tone; and the correlation between the contrastive tones of Yalálag Zapotec and their pattern of lateralization.

2 **Phonological preliminaries.** Yalálag Zapotec (henceforth YZ) is an Otomanguean language spoken in the district of Villa Hidalgo, municipality of Villa Alta, Oaxaca, by 5,000 speakers. An undetermined number of YZ speakers have also settled in Los Angeles. Yalálag Zapotec has three lexical tones: high (H), low (L) and falling (F). In what follows, I will describe the basics of YZ tone patterns. Figure 1 illustrates the f0 contours of a representative triple contrast between High, Low and Falling tones. The High and Low tones illustrated by /yá/ 'sweathouse' and /yà/ 'bell', respectively, are fairly steady, in contrast with the significant falling trajectory observable in /yâ/ 'cane'.

(1) Contrast of High, Low and Falling tone



3 **Method.** Two experiments were designed to test pitch perception and laterality in linguistic and non-linguistic stimuli. The first experiment consisted exclusively of words whereas the second included the hummed versions of words. Selective attention was investigated by three experimental conditions: non-forced (NF), forced to the left (FL) and forced to right (FR). Six right-handed native speakers of YZ (ages 30-50; 3 females, 3 males) without antecedents of hearing problems were binaurally presented with a set of minimal pairs of words differing in tone and a set of pairs of hums differing in their f0 values. Each subject was

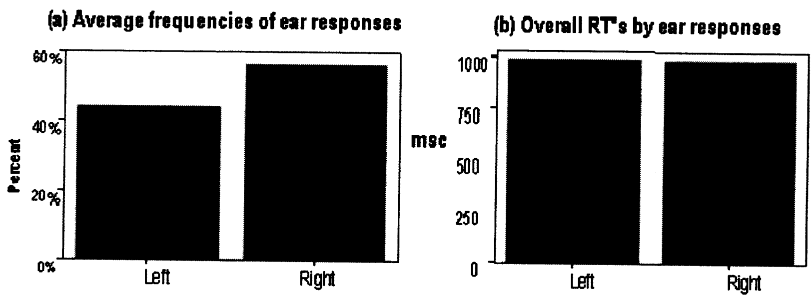
tested individually in the presence of the experimenter in the Phonetics Laboratory at UCLA or in his or her home. Trials consisted of 32 words and an equal number of hums. The experimental stimuli thus consisted of 16 dichotic pairs of words and 16 dichotic pairs of hums. Each stimuli token was presented to each ear counterbalanced, so that the total number of tokens was 32 per condition (32x3 = 96 total). Other things being equal, variation in f0 was the only difference between the members of a pair. In order to control for possible differences related to lexical accessibility, the words were of the same grammatical category and presumably of equal frequency. The same native speaker whose voice was used to produce the word stimuli also produced the hummed versions of the same words. The subjects listened to the same words (in isolation) that would be tested in a binaural condition before commencing the experimental trials. There was a practice session before the experimental trials. The words in the practice session were different from those used in the experiment. The subjects repeated the practice trials as many times as they wanted until they felt confident with the task. The subjects were asked to repeat, as soon as possible, the word or hum that they had heard. The responses were tape-recorded. The method of using oral responses was intended to reveal perceptual accuracy, since the subjects had to match one of the words of the dichotic pair. In standard procedures of dichotic listening tasks, the subject is asked to press a button which selects one of a pair of words or images from a computer screen. However, there is evidence showing that handedness has some effect biasing such responses (see Mazzucchi *et al.* 1981 for discussion on the relevance of this factor). The present design, using oral responses, aimed to avoid a handedness bias, especially because the interference of reading with the perception of dichotic stimuli is not yet fully understood. Moreover, since YZ is a language without a written tradition, the protocol in which a written word is selected from a screen would have been clearly inappropriate. The experiment lasted between 90 and 120 minutes. There was a recess of five to ten minutes between each section of the experiment. Data were collected, recorded and processed with Psyscope and independent recordings of the sessions were made on analogue tapes. Statistical analysis were obtained with SPSS software. The dichotic stimuli were normalized in duration and amplitude (70 dB) so that, other things being equal, the only difference between the members of each dichotic pair was its fundamental frequency. Stimuli normalization was done with the resynthesis function of the Praat program (Boersma and Weenick 2002). In order to ascertain the accuracy of the responses after the experimental session was over, each subject listened to his/her responses and gave the corresponding glosses in Spanish or English to the experimenter. In the case of hums, the f0 traces of the oral responses recorded were inspected to confirm a match with one of the components of the members of the dichotic stimulus.

4 Hypotheses. Hypothesis 1. Based on the robust evidence accumulated over the last decades, the basic hypothesis to test in this study is that if lexical tone entails the processing of linguistic stimuli, then a right ear advantage (REA) is predicted, whereas a left ear advantage (LEA) would be expected for the processing of non-linguistic stimuli such as hums. Hypothesis 2. If selective attention is conceived as a condition that enhances responses from the attended ear, it is expected that the responses of the attended ear (left or right) will increase according to the forced condition, in comparison with the non-forced condition. Thus, specifically, a greater REA is predicted for the lexical tone stimuli when the condition is forced to the right ear than in the non-forced condition. On the other hand, a reduction in responses from the right ear is expected when the condition is forced to the left. The opposite tendency is expected for the non-linguistic stimuli; that is, a greater LEA is predicted when the condition is forced to the left in comparison to the non-forced condition. If the results show the reverse tendency, this would suggest that selective attention is not a process of enhancement of the attended ear but rather a process that inhibits the intrusions from the non-attended ear. Hypothesis 3. A REA is expected for the processing of lexical tone. However, it is not immediately obvious what would be the pattern of lateralization among the three different tones. Nevertheless, it is possible to advance two potential hypotheses with respect to the saliency of the tones in a dichotic situation based on the psychoacoustic properties of the stimulus: the complexity hypothesis and the excitatory hypothesis. In the complexity hypothesis, the degree of complexity of the signal will resolve the decision of subjects regarding the perception of the dichotic stimuli. Thus, because the falling tone is a more complex signal, it could be considered inherently more distinguishable than a simple level tone, low or high and therefore it is expected a REA for falling tones. In the excitatory hypothesis, the characteristic frequencies of excitation of each tone will be computed to decide the perception of the tone in a dichotic condition. There is evidence that the excitatory patterns of high frequencies over the auditory nerve fibers are 'stronger' than those of low frequencies. Thus, the excitatory hypothesis predicts that in the dichotic pair high-low, the high tone will have perceptual saliency over a low tone. The two hypotheses can be empirically tested in the present study since in the duplet high-falling tones, both hypotheses can be valid. Thus, if the results indicate a greater advantage of H tone over F tone, we can propose that excitatory 'saliency' overrides complexity of the signal. However, if Falling has a greater advantage over H, that would indicate that complexity overrides saliency.

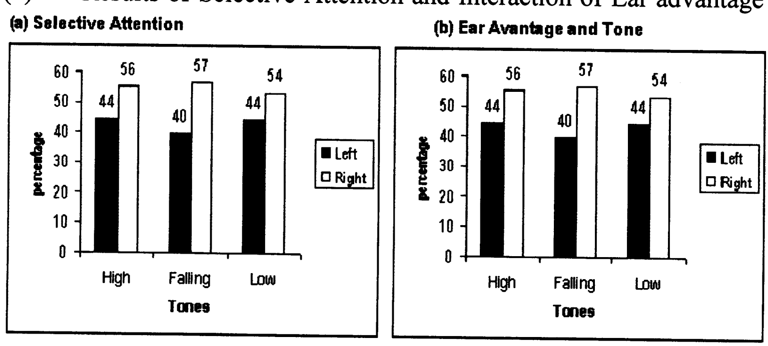
5 **Results.** Linguistic stimuli. Overall, the combined results across the three conditions (NF, FR and FL) showed a tendency towards responses from the right ear in processing lexical tone (see Figure 2a): Right ear 56% (316) versus left ear 44% (248). The same tendency was observed across all speakers, except one, EC, who showed the opposite tendency. The results showed an essentially identical reaction time for responses coming from the left ear and those from the right ear (mean 990 ms versus 985 ms (F .762, df 2, sig. 467)). Overall, the results hold across speaker. Examination of the results for selective attention (presented in Figure (3a) below) revealed an overall REA, yet this advantage showed a higher score when the attention was non-forced or when the attention was forced to the right than when attention was forced to the left. The same tendency was observed across speakers (except for EC, who showed the opposite trend). When attention was forced to the right, an increased REA was observed, compared to non-forced and forced to the left conditions. The findings regarding the interaction of tone

and ear advantage are summarized in Figure (3b). The results showed a robust trend for a REA for all the tones. Nevertheless, a closer inspection of the data revealed intra-listener variation. The results for each subject are presented in Figure (4) below. First, there was a robust REA for falling tone across four speakers (except EC). Second, there was a robust REA for low tone across speakers (moderate in EV and LEA for EC). Third, there was a robust REA for high tone in three speakers (SA, EX, RR); no ear advantage in two (JB and EV); and LEA for one (EC).

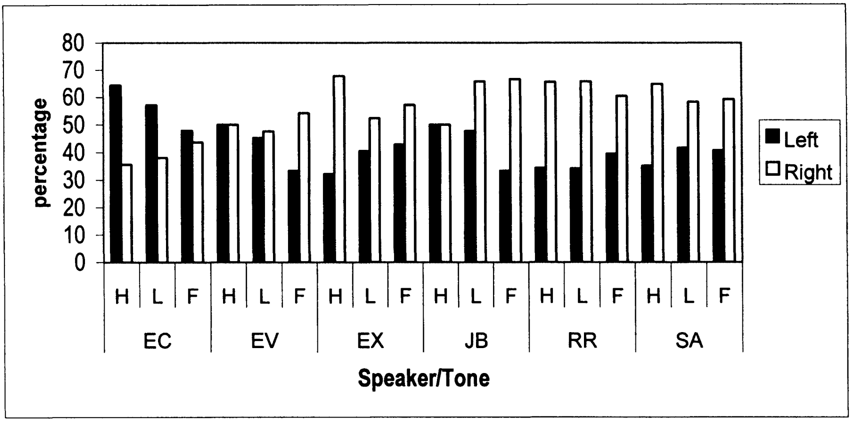
(2) Average frequencies and reaction time results



(3) Results of Selective Attention and Interaction of Ear advantage and Tone

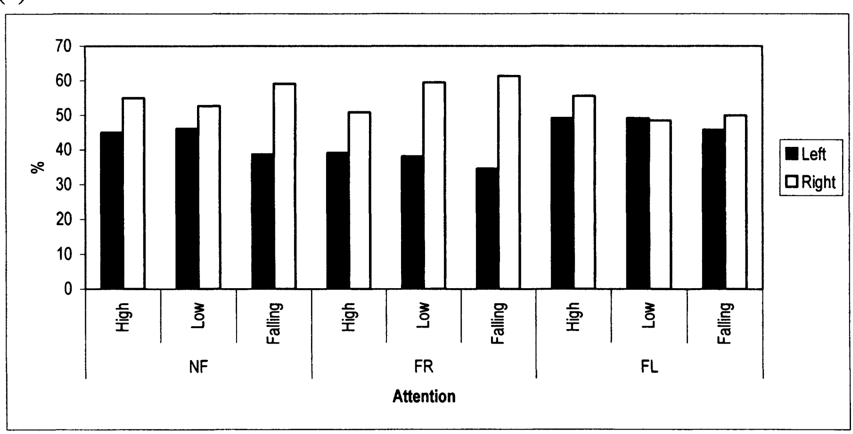


(4) Intralistener variation by tone and ear advantage



With respect to the interaction of tone and selective attention, the results showed the following patterns. First, there was a REA for all the tones except low in the forced to the left condition (FL). Second, high and falling tones showed a REA regardless of the attention condition. Third, with attention forced to the left, falling and low tones showed a greater incidence of intrusions from the left ear than did high tone. The results regarding the interaction of dichotic pairs of tone and selective attention are summarized in Table 6. In general, high and falling tones showed a consistent REA. Selective attention to the right enhanced the overall REA, although the highest differences were found in NF. Low tone showed the 'weakest' score when matched with H or F, as indicated by the absence of a clear EA in the pair F-L, and the LEA in the dichotic pair H-L.

(5) Selective attention and tone



(6) Dichotic pairs of tone and selective attention interaction

Tone Left-Tone Right	Overall	FL	NF	FR
H-F	REA (65-35%)	REA (64-34%)	REA (64-36%)	REA (67-33%)
F-H	REA (64-36%)	no EA (50-50%)	REA (75-25%)	REA (66-34)
H-L	LEA (51-49%)	no EA (50-50%)	LEA (63-37%)	REA (61-39%)
L-H	REA (58-42%)	LEA (49-51%)	REA (66-34%)	REA (59-41%)
L-F	REA (60-35%)	REA (50-44%)	REA (68-26%)	REA (61-35%)
F-L	no EA (48-47%)	LEA (50-44%)	REA (50-47%)	REA (50-44%)

Non-linguistic stimuli. Overall, the results showed a greater incidence of responses from the left ear in processing hums: 52% (151) > 48% (137). Only one speaker (SA) deviated from this tendency, showing a slight preference for the right ear (52% > 48%). With respect to the interaction of selective attention and the ear response the results showed a greater score for the left ear in both forced conditions than in the non-forced condition. Although the scores for the forced conditions are similar, the individual scores differ with respect to the distribution of responses. Table 7 summarizes these findings. The results regarding the interaction of tone type and response showed that listeners were more likely to identify the tones when they were delivered to the left ear than when they were delivered to the right ear: High= 50.5% > 49.4%, Falling= 53% > 47% and Low = 53.5% > 46.5 %.

(7) Ear responses and forced conditions for hums

	Forced to the left	Forced to the right	Non-forced
Left ear	55%	55%	49%
Right ear	45%	45%	51%

The results for the specific interactions of tones in the dichotic task are summarized in Table (8). Conspicuously, there is no homogeneous distribution of tone identification. However, a few generalizations emerge. First, low tone was less often identified than high or falling tone in the dichotic stimulus. This tendency was observed regardless of the condition directing the attention. This is absolute in the dichotic pairs involving high and low tone. As for the pairs involving falling and low tone, the tendency is similar with the exception of the pair L-F under the condition forcing attention to the right ear and the non-forced condition, where there was no ear advantage at all. Second, the identification of pairs formed by high and falling tone performed below chance when the high tone was delivered to the right ear. The results showed a non-significant interaction between reaction time and ear response (one factor ANOVA $F(1, 289) = 0.218$ $p > .05$). No other significant interactions between any of the conditions tested were observed.

(8) Ear advantage of dichotic pairs and attention conditions for hums

	Overall	FL	NF	FR
H-F	REA 39%-61%	REA 33%-67%	REA 33%-67%	no EA 50%-50%
F-H	no EA 50%-50%	no EA 50%-50%	no EA 50%-50%	no EA 50%-50%
H-L	LEA 57%-43%	LEA 58%-42%	LEA 54%-46%	LEA 58%-42%
L-H	REA 47%-53%	REA 46%-54%	REA 46%-54%	REA 42%-58%
L-F	REA 46%-54%	REA 33%-67%	no EA 50%-50%	LEA 55.5%-44.5%
F-L	LEA 65%-35%	LEA 77%-23%	no EA 50%-50%	LEA 66%-34%

6 **Summary and Discussion.** The results of the present study have shown that there is a consistent REA in the processing of lexical tone, suggesting a left hemisphere advantage for the processing of tone by listeners of Yalálag Zapotec. In contrast, there was a tendency for a LEA for hums, which could be interpreted as a right lateralization for non-linguistic stimuli. On this interpretation the findings of the present study are consistent with the theory that proposes that the left hemisphere is highly specialized for linguistic functions, whereas the right hemisphere is less involved in linguistic tasks (Fitch *et al.* 1997, Moffat and Hampson 2000, among others). Particularly, the results of this study show a consistent REA for phonemic tones; this suggests that there is a left hemisphere dominance in the processing of f0 when it is part of the linguistic system. The results obtained in the present study are consistent with similar findings by Van Lancker and Fromkin (1973, 1978), Wang *et al.* (2001), Moen (1993), Gandour *et al.* (1988) and Hugdahl *et al.* (1999), all of which report a REA in the processing of phonemic tone. Likewise, the findings of the present study indicating a right hemisphere dominance for processing f0 in a non-linguistic context also confirm previous research (Blumstein and Cooper, 1974, Goodglass and Calderon 1977, Perkins *et al.* 1996, Ross *et al.* 1988). Two general accounts have been proposed regarding the hemispheric specialization of pitch: a theory of so-called "functional hierarchy" of pitch processing, put forward by Van Lancker (1980) and a theory of pure pitch processing, represented mainly by Ivry and Robertson (1998). In Van Lancker's account the association of pitch processing with the left or right hemisphere depends on its function. Hence, more specialized linguistic functions of pitch, as in the use of phonemic tone, would be lateralized to the left hemisphere, whereas less or non-linguistic uses, such as pure tones, singing or even intonation, would be processed primarily in the right hemisphere. According

to Ivry and Robertson, the asymmetries observed in linguistic tasks (as well as a broader range of perceptual phenomena) are explained as a dissimilar "power" of each hemisphere, which is further conditioned by attentional mechanisms. Overall, the findings obtained in this study are consistent with the view that proposes a functional processing of pitch. Thus, since phonemic tone is a highly specialized linguistic function in the sound pattern of YZ, it is processed primarily in the left hemisphere. In contrast, because of the non-linguistic nature of hummed words, these tended to be processed by the right hemisphere. It is not clear how Ivry and Robertson's proposal could account for these facts and the asymmetrical results based on the type of stimuli under different attention conditions. The results in the present study indicate a consistent tendency for a REA in processing lexical tone; however, the amount of responses from the left ear suggests that the right hemisphere is also active in the processing of tone. In fact, recent studies suggest that this is the case. Andy and Bhatnagar (1984) presented electroencephalographic evidence of direct brain stimulation showing that the right hemisphere is active in performing linguistic tasks. Other studies on split-brain subjects have also indicated that the right hemisphere is active in the processing of some linguistic functions (Gazzaniga 2000, Zaidel 1976). Thus, the results of the present study are consistent with the view suggesting that although the left hemisphere is prepotent for linguistic functions, the right hemisphere makes a significant contribution to the entire linguistic processing in ways yet undefined. The specific nature and extent of its contribution should be further investigated. This study also investigated attentional effects in dichotic listening to pitch pairs. Traditional models of dichotic listening would have predicted a null effect of attention on ear advantage (e.g. Kimura 1967). Instead, the findings obtained in the present study partially confirmed the initial hypothesis anticipating a positive interaction between ear advantage and directed attention of the subjects. The results are thus consistent with several recent studies addressing the issue of selective attention in dichotic listening (Hugdahl and Anderson 1986, Asbjørnsen and Hugdahl 1995). However, against the predictions about the lateralization for the non-linguistic stimuli, the results showed a LEA in the forced conditions and a minor REA in the non-forced condition. This pattern can be interpreted as an effect of attention enhancing the right-hemisphere processing of non-linguistic stimuli. Thus, the findings on the effects of directed attention in dichotic listening to pitch stimuli support the hypothesis that forcing attention to one of the ears enhances the performance of the attended ear rather than suppressing intrusions from the non-attended ear. One of the novel issues addressed in this study was the relationship between the individual pitches in a dichotic listening condition. The results showed that, unlike the linguistic stimuli, the interaction of non-linguistic pitch stimuli did not show a pattern at all. Such a result may suggest that the non-linguistic nature of hums is either irrelevant or more demanding for processing. The predictions concerning the interaction among individual lexical tones were supported in part. First, it was hypothesized that the low tone would be the least salient tone in the dichotic task. The overall results for both types of stimuli confirmed this expectation, since when low tone was delivered to the right ear, showed more intrusions from the left ear. The vulnerability of low tone was also observed when selective attention was considered. Even though, in the linguistic stimuli, there was a REA for low tone delivered to the right ear under forced to the right condition, the magnitude of the enhancement was smaller than those obtained for high and falling tones. Second, I suggested two possibilities for the interaction between H and F tones. In one of them it was predicted that the inherent complexity of F would be more salient than the characteristics defining H. In the other, the high frequencies of H would exceed the perceptual properties of F. If the non-forced condition is taken as a diagnostic of the natural trend of lateralization, the higher score for high tone in the dichotic pair F-H (where high is delivered to the right ear) may indicate a perceptual preference for high frequencies over the complex signal of F. Nevertheless, if the condition forcing the attention to the left is considered as an index to measure intrusions to the prepotent left hemisphere, then the crucial results showing a REA in the pair H-F and lack of ear advantage in the pair F-H may indicate that the perceptibility of F is greater than that of H. Thus far, the conclusion seems to be that there is no single factor determining the perception of dichotic stimuli; rather both factors, the complexity of the signal and the patterns of excitation appear to be involved.

7 **Concluding remarks.** Based on the index of ear advantage, the evidence suggests that the processing of lexical tone by Yalálag Zapotec listeners occurs in the left hemisphere, in contrast with non-linguistic stimuli, which are processed in the right hemisphere. The results are consistent with other similar studies of tone languages (Moen 1993, Van Lancker and Fromkin 1973, 1978, Wang *et al.* 2001). Furthermore, the results of the present study should be evaluated in the light of two recent PET studies (Gandour *et al.* 2000, Klein *et al.* 2001) investigating the perception of f0 in tone languages using a similar behavioral paradigm to the present study have found a significant activation in frontal, parietal and parieto-occipital regions of the left hemisphere only in native listeners of tone languages when presented to linguistic stimuli, in contrast with listeners of non-tone languages who showed a right inferior frontal cortex activation. Since this study has presented initial evidence to suggest a pattern of brain lateralization for pitch processing in Yalálag Zapotec, it would be desirable to examine the current findings in the light of modern neuroimaging techniques to fully understand the hemispheric specialization of pitch in a tone language in future research.

References

Andy, Orlando J. and Subashi C. Bhatnagar. 1984. Right-Hemispheric language evidence from cortical stimulation. *Brain and Language* 23, 159-166.

Asbjørnsen, A. and K. Hugdahl, 1995. Attentional effects in dichotic listening, *Brain and Language* 49, pp. 189-201.

Blumstein, S., and W. Cooper .1974. Hemispheric processing of intonation contours. *Cortex,* 10, 146–158.

Boersma, Paul and David Weenick. 2002. *Praat. Doing phonetics by computer.* Version 4.0.11 Institute of Phonetic Sciences of the University Amsterdam.

Corina, David. 1998. Aphasia in signed languages. *Aphasia in atypical populations.* Patrick Coppens, Yvan Lebrun and Anna Basso (eds.) Lawrence Erlbaum Associates, Publishers. 261-309.

Fitch R. H., S. Miller, and P. Tallal. 1997. Neurobiology of speech perception. *Annual Review of Neuroscience,* 20, 331–353.

Gazzaniga, M. S. 2000. Cerebral specialization and interhemispheric communication. Does the corpus callosum enable the human condition? *Brain,* 123, 1293-1326.

Gandour, J., S. H. Petty and R. Dardarananda. 1988. Perception and Production of Tone in Aphasia. *Brain and Language*, 35, 201-240.

Gandour, J., D. Wong, L. Hsieh, B. Weinzapfel, D. Van Lancker and G. Hutchins. 2000. A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12.207-22.

Goodglass, Harold and M. Calderon. 1977. Parallel processing of verbal and musical stimuli in right and left hemispheres. *Neuropsychologia,* 15, 397-407.

Hugdahl, Kenneth (ed.) 1988. Handbook of dichotic listening : theory, methods, and research. Chichester: New York.

Hugdahl, Kenneth; Kolbjorn Bronnick; Soren Kyllingsbae; Ian Law; Anders Gade and Olaf B. Paulson. 1999. Brain Activation during Dichotic Presentations of Consonant-Vowel and Musical Instrument Stimuli: A 15O-PET Study. *Neuropsychologia*, 37, 4, Apr, 431-440.

Hughdal, Kenneth and L. Anderson 1986. The 'forced attention paradigm' in dichotic listening to CV-syllables: A comparison between adults and children. *Cortex.* 22, 417-432.

Ivry, Richard and Lynn Robertson. 1998. *The two sides of perception.* MIT Press.

Kimura, D. 1961. Cerebral Dominance and the Perception of Verbal Stimuli. *Canadian Journal of Psuchology*, 15: 166-171.

Kimura, D. 1967. Functional Asymmetry of the Brain in Dichotic Listening. *Cortex*, 3:163-178.

Klein, D., R. J. Zatorre, B. Milner, V Zhao. 2001. A Cross-Linguistic PET Study of Tone Perception in Mandarin Chinese and English Speakers. *NeuroImage*, 13.646-53.

Mazzucchi, A., M. Parma and R. Cattelani. 1981. Hemispheric dominance in the perception of tonal sequences in relation to sex, musical competence and handedness. *Cortex,* 17, 291–302.

Moen, Inger. 1993. Functional Lateralization of the Perception of Norwegian Word Tones – Evidence from a Dichotic Listening Experiment. *Brain and Language*, 44, 4, May, 400-413.

Moffat, Scott D and Elizabeth Hampson. 2000. Language Lateralization?. *Neuropsychology*, 14, 1, Jan, 71-81.

Penfield, Wilder and Lamar Roberts. 1959. *Speech and brain-mechanisms.* Princeton, N.J., Princeton University Press.

Perkins, J. M. Baran, and J. A. Gandour. 1996. Hemispheric specialization on processing intonation contours. *Aphasiology*, 10, 343-362.

Ross, Elliott D., Jerold A. Edmonson, G. Burton Seibert and Richard W. Homan. 1988. Acoustic analysis of affective prosody during Right-Sided Wada Test:

A within-subjects verification of the right hemisphere's role in language. *Brain and Language*, 33, 128-145.
Van Lancker, Diana. 1980. Cerebral lateralization of pitch cues in the linguistic signal. *Papers in Linguistics: International Journal of Human Communication,* 13, 201–277.
Van Lancker, Diana and Victoria Fromkin. 1973. Hemispheric specialization for pitch and "tone": Evidence from Thai. *Journal of Phonetics,* 1, 101–109.
Van Lancker, Diana and Victoria Fromkin. 1978. Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers *Journal of Phonetics,* 6, 19–23.
Wang, Yue, Allard Jongman and Joan A. Sereno. Dichotic Perception of Mandarin Tones by Chinese and American Listeners. 2001. *Brain and Language.*
Zaidel, Eran. 1976. Language, dichotic listening, and the disconnected hemispheres. (ms)

Heriberto Avelino
Department of Linguistics, University of California at Los Angeles
3125 Campbell Hall
Los Angeles, California 90095-1543

avelino@ucla.edu

Variability and Constancy in the Articulation and Acoustics of Pima Coronals[1]

HERIBERTO AVELINO and SAHYANG KIM
Department of Linguistics, UCLA

1. Introduction

Pima is a Uto-aztecan language spoken in the state of Arizona, U.S.A. and Sonora, Mexico by 11,819 speakers. Pima presents a contrast among denti-alveolar /t̪, d̪, s/, alveolar /d/, post-alveolar /ʂ/ and alveo-palatal /tʃ, dʒ/ Less than 15 % of 452 languages distinguish two coronal stops in place of articulation, according to the UPSID corpus (Maddieson 1984, Maddieson and Precoda 1992). Since the distinction is rather uncommon across languages, our primary goal is to provide a detailed description of the phonetic properties of the sounds involved in this contrast. Based on this description, our study addresses the question about the specific characterization of these consonants. Furthermore, we discuss the classification proposed for related languages, such as Tohono O'odham and Pima Bajo (Escalante and Estrada Fernandez 1993, Saxton 1963). These studies claimed that the contrast between /d̪, s/ and /d, ʂ/ corresponds to a difference between dental and retroflex sounds. However, these studies lack the precise description about how the retroflex sounds are articulated. The term 'retroflex' generally describes an articulation made with the inferior part of the tongue by curling the tip. Nevertheless, the label has been used for quite different articulatory gestures across languages (Ladefoged and Maddieson 1996). In our earlier fieldwork on Pima, we have observed that the articulation of apical sounds might not be well described by the gesture of curling the tongue tip. The detailed articulatory study will shed a light on this issue.

[1] We would like to thank the five Pima speakers who decided to remain anonymous and collaborated in this project. We acknowledge the authorities at the Gila River Reservation for granting us the permission to do the fieldwork. Without their collaboration this work could not have been possible. We are especially grateful to Mr. Virgil Lewis who has taught us his language and helped us in the preparation for the fieldwork. We wish to thank Jenny Ladefoged and Peter Ladefoged who supervised the fieldwork and guided us through the whole process of this study. This project was supported by a grant from the Institute of American Cultures through the American Indian Center at UCLA and by the Department of Linguistics at UCLA. The first author acknowledges a UC-MEXUS/CONACYT fellowship.

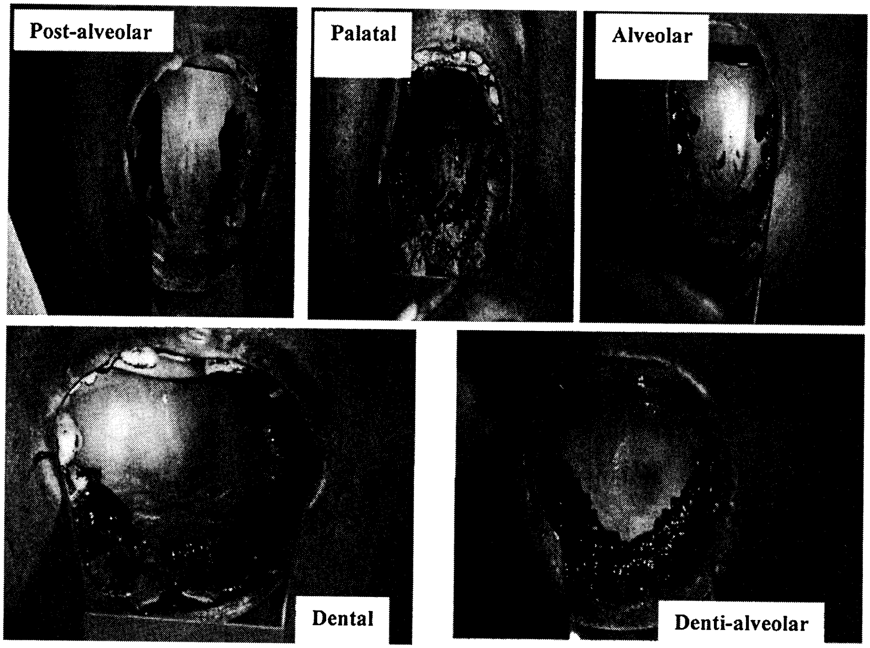
The compression of the coronal consonants in a narrow articulatory space is also relevant to phonetic theory because it allows us to observe the degree of variability and/or constancy in the production of these sounds. Some phonetic theories proposed that the sounds that are perceptually salient or typologically common can show articulatory variability to the extent that the corresponding acoustic parameters show relative constancy (Stevens 1985, 1989, 1998). Similarly, it has been suggested that small phonemic systems demand less perceptual distinctiveness than larger systems (Lindblom 1990). These ideas allow us to predict that a speech sound in a dispersed articulatory space may allow greater articulatory variability than that in a compressed space, as far as its acoustic properties remain constant. Hence, our secondary goal is to examine the correlation between the articulatory and acoustic properties of Pima coronal sounds.

2. Method

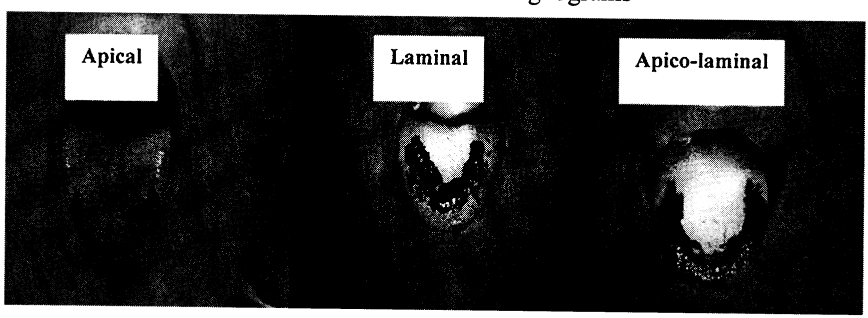
Data collection and subjects. The corpus investigated consisted of six words containing the target segments in a /...VCV.../ context (where V is [-high]). Subjects were asked to repeat each word three times. The materials were collected in fieldwork in the Gila River reservation. Three female and three male native speakers of Pima, ranging from 40-60 years of age participated in this study. Audio data were recorded using digital audio tapes and digitized at a 22,050 Hz sampling rate. Video data were recorded in a digital camera and transferred to a computer.

Articulatory Measurements. To record articulatory data, we adopted the method of static palatography (Anderson 2000, Dart 1993, Ladefoged to appear, among others). In order to obtain palatogram data, the tongue of a speaker was coated with a colored substance (an innocuous mixture of charcoal and olive oil) so that when the sound being studied was articulated, there was an imprint of the articulation on the roof of the mouth. This imprint was reflected in a mirror and captured as a video image. In linguograms, the process was reversed. The roof of the mouth was coated, and the imprint on the tongue (reflects the location of linguo-palatal contact) was recorded as a video image. In order to relate palatographic/linguographic images to the actual articulation, a life-size dental impression was made of each speaker's palate and cast in plaster. For each still frame, measurements from the vertical and horizontal axes were calibrated according to the measurements of the life-size casts for each speaker. For palatograms, data were classified as post-alveolar, palatal, dental, denti-alveolar and alveolar according to the measurements and visual inspection of the images compared to the plaster casts. Figure (1) illustrates representative palatograms of each category. For linguograms, data were classified as apical, laminal and apico-laminal according to visual inspection of the images. Figure (2) shows typical examples of the three categories.

(1) Classification of palatograms



(2) Classification of linguograms



Acoustic Measurements. Several measurements were obtained to perform acoustic analyses. We measured the centroid frequency and spectra of fricatives and affricates, formant transitions, stop burst release amplitude for stops and durational properties. The centroid frequency for fricatives were obtained from FFT spectra computed over a 25 ms window in a frequency range of 1000 Hz to 10000 Hz. Spectral properties were obtained from FFT spectra using a 1024 point frame, which amounted to 46 ms window. Formant transitions (F1, F2, F3, F4) of the surrounding vowels were obtained. Vowel formants were measured at the offset of the preceding vowel and the onset of the following vowel. The formants at the steady state of vowels were also measured as a reference point. Relative 'vowel-burst' amplitude was calculated from the difference between the maximum amplitude of the vowel and the peak stop burst amplitude. Various analyses of durational properties were obtained fricative duration, voice onset time, stop closure duration and rate of voicing during the acoustic closure were measured.

3. Results

3.1. Articulation

Stops. The results regarding the production of stops are summarized in Table (3). /t̪/ and /d̪/ were articulated as dental or denti-alveolar and laminal or apico-laminal. /d/ was articulated as alveolar or post-alveolar; most speakers articulated /d/ as apical, except for the two who pronounced it with the tongue blade.

(3) Stop articulation

		SPK 1 (f)	SPK 2 (f)	SPK 3 (f)	SPK 4 (m)	SPK 5 (m)	SPK 6 (m)
/t̪/	Palato-gram	Denti-alveolar	Denti-alveolar	Dental	Denti-alveolar	Denti-alveolar	Denti-alveolar
	Linguo-gram	Laminal	Apico-laminal	Laminal	Apico-laminal	Laminal	Apico-laminal
/d̪/	Palato-gram	Dental	Denti-alveolar	Denti-alveolar	Denti-alveolar	Denti-alveolar	Denti-alveolar
	Linguo-gram	Laminal	Laminal	Laminal	Laminal	Laminal	Apico-laminal
/d/	Palato-gram	Post-alveolar	Alveolar	Alveolar	Alveolar	Post-alveolar	Alveolar
	Linguo-gram	Apical	Apical	Apical	Laminal	Laminal	Apical

Fricatives. /s/ was articulated as denti-alveolar or alveolar and laminal whereas /s̠/ was consistently post-alveolar and it was produced with the tongue tip for most of the speakers, although two speakers produced it with the blade. The channel width was consistently wider in /s̠/ than in /s/, as shown in Table (4).

(4) Fricative articulation

	/s/		/s̠/	
	Palatogram	**Linguogram channel width (cm)**	**Palatogram**	**Linguogram channel width (cm)**
SPK1 (f)	Denti-alveolar	Laminal (0.8)	Post-alveolar	Laminal (0.7)
SPK2 (f)	Alveolar	Laminal (0.4)	Post-alveolar	Apical (0.8)
SPK3 (f)	Denti-alveolar	Laminal (0.7)	Post-alveolar	Laminal (0.8)
SPK4 (m)	Alveolar	Laminal (n/a)	Post-alveolar	Apical (n/a)
SPK5 (m)	Denti-alveolar	Laminal (0.6)	Post-alveolar	Apical (0.95)
SPK6 (m)	Denti-alveolar	Apico-laminal (0.6)	Post-alveolar	Apical (0.8)

Affricates. As seen in Table (5) /tʃ/ was consistently produced as laminal. It showed wider variation in place of articulation, ranging from denti-alveolar to palatal. /dʒ/ was also consistently produced as laminal. It also showed a great deal of variation in place of articulation, ranging from alveolar to palatal.

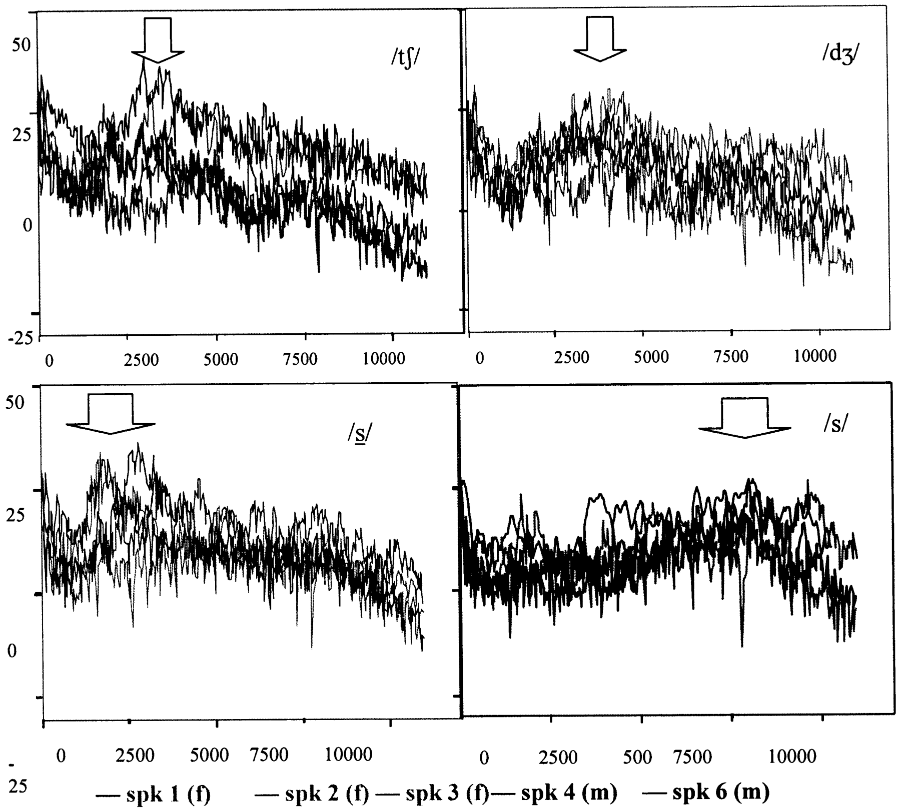
(5) Affricate articulation

	/tʃ/		/dʒ/	
	Palatogram	Linguogram	Palatogram	Linguogram
SPK1 (f)	Alveolar	Laminal	Alveolar	Laminal
SPK2 (f)	Post-alveolar	Laminal	Post-alveolar	Laminal
SPK3 (f)	Post-alveolar	Laminal	Post-alveolar	Laminal
SPK4 (m)	Palatal	Laminal	Palatal	Laminal
SPK5 (m)	Alveolar	Laminal	Alveolar	Laminal
SPK6 (m)	Denti-alveolar	Laminal	Alveolar	Laminal

3.2 Acoustics

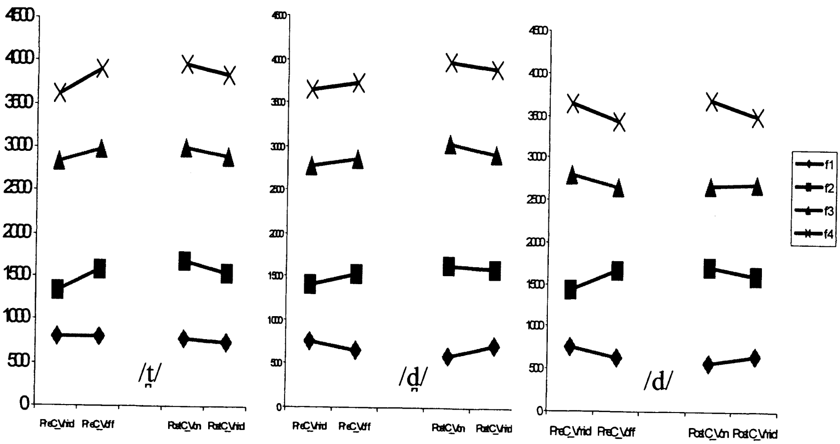
Spectral Properties. Affricates /tʃ/ and /dʒ/ showed the greatest amount of amplitude between 2500 and 4000 Hz. The amplitude dropped off in the higher frequencies. The spectrum of /s̠/ showed the greatest amount of amplitude around 2500 Hz, which went down in the higher frequencies. However, the spectrum of /s/ differed from the others. The amplitude peaks were in the frequencies after 7000 Hz, except for one speaker whose spectrum of /s/ showed more or less flat pattern. Figures in (6) show the results.

(6) Spectral properties of fricatives and affricates



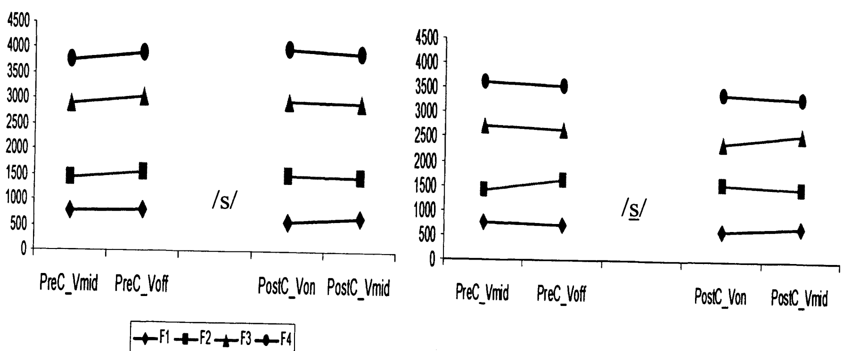
Formant transitions. Formant values at the vowel midpoints differed according to the subtle articulatory differences of the surrounding stops: vowels following dental /t̪/ and /d̪/ showed higher F3 and F4 than those following alveolar /d/, and vowels preceding alveolar /d/ showed higher F2 than those preceding dentals /t̪/ and /d̪/. As shown in Figure (7), these differences further led to different trajectories in formant transitions. Dental /t̪/ and /d̪/ raised F3 and F4 of the preceding and following vowels, but alveolar /d/ lowered F3 of the preceding and following vowels.

(7) Formant transitions of vowels surrounded by stops



Regarding the contrasts in fricatives, our results showed that denti-alveolar /s/ raised the F3 and F4 of the preceding and following vowels, whereas post-alveolar /s̠/ lowered the F3 of the preceding and following vowels. Post-alveolar /s̠/ lowered the F4 of the preceding vowel as well. Vowels that preceded denti-alveolar /s/ showed lower F2 than those which preceded post-alveolar /s̠/, as shown in Figure (8).

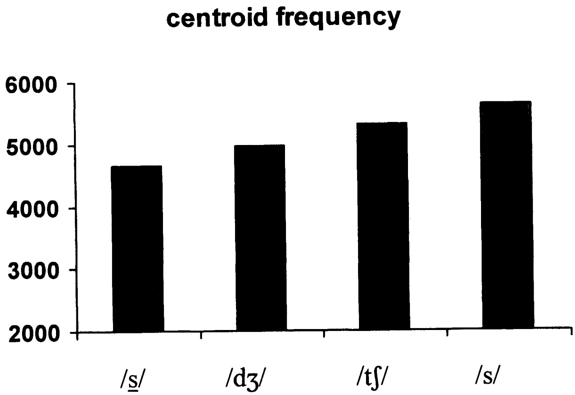
(8) Formant transitions of vowels surrounded by fricatives



The observed formant transition patterns of /tʃ/ and /dʒ/ were similar to that of /s̠/, in that they both lowered the F3 of the preceding and following vowels.

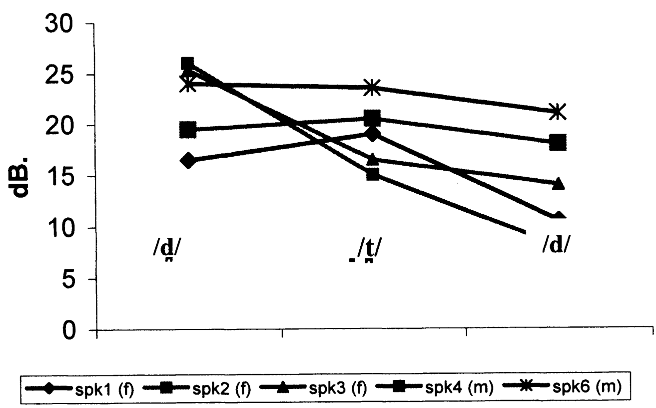
Centroid frequency for fricatives. The centroid frequency of /s/ was higher than those of /s̠/, /tʃ/ and /dʒ/ as in Figure (9). A *t*-test result showed that the difference between /s/ and /s̠/ was statistically significant ($p < .001$).

(9)



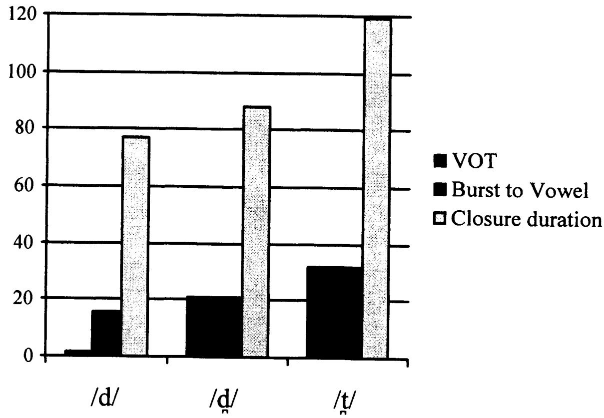
Amplitude difference between vowel maxima and stop burst. Vowel-burst differences were smaller (higher amplitude burst) for alveolar /d/ than for dentals /d̪/and /t̪/. An overall significant difference was found between /d̪/and /d/ (p= .0002), /t̪/ and /d/ (p=.03) but not between /d̪/ and /t̪/.

(10) Amplitude difference between vowel maxima and stop burst



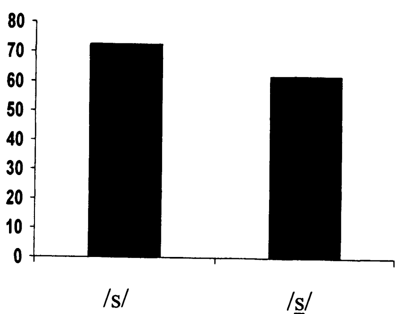
Stop durations. The mean burst to vowel duration value of /d/ (15.6 ms) was shorter than that of /d̪/ (20.9 ms) and /t̪/ (32.1 ms). The difference /d/ and /t̪/ was significant. The mean VOT value of /d/ (1.8 ms) was shorter than that of /d̪/ (20.9 ms) and /t̪/ (32.1 ms). There was a significant difference between /d/ and /t̪/. The mean closure duration value of /d/ (77.2 ms) was shorter than that of /d̪/ (87.7 ms) and /t̪/ (119.5 ms). There was a significant difference between /d/ and /t̪/, and between /d̪/ and /t̪/. The results are summarized in (11).

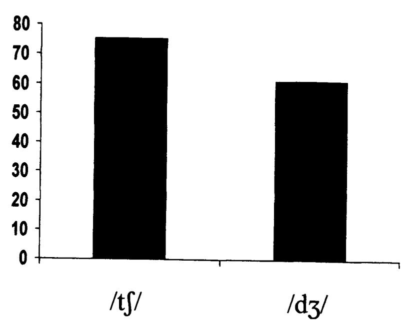
(11) Acoustic durations for stop consonants



Fricative durations. The duration of /s/ was longer than that of /s̱/, but the difference was not significant. The closure duration for /tʃ/ was significantly longer than that of /dʒ/ (*p* <.05). There was no significant difference in the duration of frication between /tʃ/ and /dʒ/. The results are summarized in (12).

(12) **fricative duration** **closure duration for affricates**





4. Discussion

The phonetic description provided in this study showed that /t̪/, /d̪/, and /s/, the sounds which were previously categorized as dentals, are produced with dental or denti-alveolar constriction using the tongue blade. In contrast, /d/ was articulated with the tip or the blade of the tongue against the alveolar ridge or in the post-alveolar region. /s̱/ had its point of constriction in the post-alveolar region and the tongue contact was made either by the tip or the blade. Palatal sounds

/tʃ/ and /dʒ/ were produced with the tongue blade, and their places of articulation ranged from denti-alveolar to palatal.

Previous descriptions of O'odham languages (Saxton 1963) considered /d/ and /s̱/ as retroflex sounds. Our results indicate that a more accurate account of these sounds should describe them as retracted articulations, which, at the phonetic level, should be differentiated from the retroflex sounds made by curling the tongue tip back (Ladefoged and Maddieson 1996).

The results showed a correlation between the place of articulation and the tongue contact. In stops, dentals and denti-alveolars involved laminal contact, whereas alveolars tended to involve apical contact. In fricatives, alveolar fricatives involved laminal contact, and post-alveolars tended to involve apical contact. The generalization emerging from these observations is consistent with the cross-linguistic patterns observed by Ladefoged and Maddieson (1996): fronter points of constriction involve tongue blade contact and more posterior points of constriction involve tongue tip contact.

The results also indicated that there is a correlation between place of articulation and formant transitions of the surrounding vowels. Alveolar /d/ lowers F3 into the transition of the following vowel. It is due to the fact that the larger front cavity formed when the tongue tip is raised and retracted produces smaller F3 and F4. Post-alveolar /s̱/ produced a dramatic lowering of F3 of the adjacent vowel transitions, as opposed to /s/ that triggered a slight raising on the adjacent vowel formant transitions. Our findings were consistent with previous analyses of coronal consonants (Anderson 2000, Dart 1993, Gordon *et al.* to appear, Stevens 1998) suggesting that the lowering effect of /s̱/ is concomitant to the larger sublingual resonance of the consonant.

Our findings showed a correlation between burst amplitude and the active articulator. The burst amplitude was lower for laminal stops (/t̪/, /d̪/) than apicals (/d/) due to the prolonged obtsruction of the oral cavity caused by the gradual release of the tongue. This observation is consistent with the reports of previous studies of apicality in Malayalam (Jongman et al. 1985) and in Tiwi (Anderson and Maddieson 1994).

The results showed a robust correlation between the place of articulation and the energy distribution of fricatives. That is, when the front cavity length was smaller, both centroid frequency and locus of spectral noise were concentrated at higher frequencies (Stevens 1998).

We also found a strong correspondence between articulatory variation and spectral shape constancy of frication noise. The results showed that regardless of the intra-speaker variation found in the articulation of fricatives and affricates, its acoustic properties remained constant. This observation may indicate that acoustic stability of the location of amplitude peak is more important for the overall perception.

Overall, our results seem to suggest that even a language like Pima, which has phonological contrasts within a limited articulatory space, does not impose strict demands on articulatory precision to the extent that the target acoustic signal is produced.

References

Anderson, Victoria. 2000. *Giving weight to phonetic principles:The case of Place of Articulation in Western Arrernte*. PhD dissertation, UC Los Angeles.

Anderson, Victoria and Ian Maddieson. 1994. Acoustic characteristics of Tiwi coronal stops. *UCLA Working Papers in Phonetics* 87: 131-162.

Dart, Sarah. 1993. Phonetic properties of O'odham stop and fricative contrasts. *International Journal of American Linguistics* 59: 16-37.

Gordon, Matt, Paul Barthmaier and Kathy Sands. to appear. A cross-linguistic acoustic study of fricatives. *Journal of the International Phonetic Association.*

Escalante, Roberto, Zarina Estrada Fernandez. 1993. *Textos y gramatica del Pima Bajo*. Universidad de Sonora.

Jongman, Allard, Sheila E. Blumstein, and Aiti Lahiri. 1985. Acoustic properties for dental and apical alveolar stop consonants: a cross-language study. *Journal of Phonetics* 13: 235-251.

Ladefoged, Peter. to appear. *Phonetic Data Analysis: An introduction to phonetic fieldwork and instrumental techniques.* Blackwell Publishers.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's languages.* Cambridge, MA: Blackwell Publishers.

Lindblom, Bjorn. 1990. Explaining phonetic variation: A sketch of the H and H theory. In W.J. Hardcastle & A. Marchal (eds.) *Speech Production and Speech Modeling*. Dordrecht:Kluwer.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian and Kristine Precoda. 1992. *UCLA Phonological Segment Inventory Database*, UC Los Angeles.

Saxton, Dean. 1963. Papago phonemes. *International Journal of American Linguistics* 29: 29-35.

Stevens, Kenneth. 1985. Evidence for the role of acoustic boundaries in the perception of speech sounds. In V.A. Fromkin (ed.) *Phonetic linguistics: Essays in honor of Peter Ladefoged*, 243-255. London: Academic Press.

Stevens, Kenneth. 1989. On the quantal nature of speech. *Journal of Phonetics* 17: 3-45

Stevens, Kenneth. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Heriberto Avelino
Dept. of Linguistics, University of California, Los Angeles
3125 Campbell Hall
Box 951543
Los Angeles, CA 90095-1543

avelino@ucla.edu

Another Look at Australia as a Linguistic Area

CLAIRE BOWERN
Harvard University

1. Introduction

Australia has been described as an ancient and complex linguistic area, with few recoverable genetic groups (for example Dixon 1997, 2001, 2002). Evidence cited in support of the areal model of Australian linguistic relationships includes: the lack of bunching isoglosses for major typological features such as prefixation, bound pronouns, noun classes, and phonological changes; the difficulty of classifying the Pama-Nyungan daughter languages into discrete subgroups; the apparent time-depth of settlement of the continent (at least 50,000 years); and longstanding, widespread multilingualism.

There are, indeed, many isoglosses that cross-cross the continent, and progress has been slow in establishing subgrouping and completing appropriate reconstructions. However, I argue here that contrary to recent high-profile work on Australian languages, Australia is not a 'special case' where traditional methods are of no use. Moreover, I hope also to put aside the idea that the use of the terms 'genetic relationship' or 'subgroup' necessarily implies the ability to model the languages on a neat family tree. Finally, our lack of ability to model a given group of languages on a family tree does not necessarily imply a great time depth to the group.

In this paper I begin by giving a brief overview of the assumptions made in the punctuated equilibrium model. In Section 3 I summarize the arguments given to account for why more traditional models are assumed not to work in Australia. In Section 4 I summarize some of the major problems in applying punctuated equilibrium in Australia, and finally in Section 5 I present an alternative scenario which account for the facts and the patterns we see.[1]

2. Overview of punctuated equilibrium

In this section I concentrate on the major features of punctuated equilibrium as a model of language change. I will discuss the notions of 'punctuation' and

[1] This is a summary version of a much longer paper in progress, which will be included in a book on linguistic areas edited by Yaron Matras, April McMahon, and Nigel Vincent.

'equilibrium' and the problems with applying these notions to language change, and linking language change to non-linguistic events in a systematic way. I will address Dixon's criticisms of the comparative method and the family tree model, and the 'special situation' of Australian languages.

2.1 General Summary

Dixon adapted the term 'punctuated equilibrium' from evolutionary biology (and specifically Eldredge and Gould 1972) to refer to his model in which language change can occur at vastly different rates over different periods of time.

In the model, language groups are assumed to have coexisted for most of human history without a great deal of disruption. During this time, languages continue to change, largely, Dixon (1997:71) implies, by borrowing from one another. In this way traits can diffuse across an area. Thus during equilibria local as well as larger linguistic areas are created, across boundaries of genetic families. The long periods of equilibrium are interspersed with much shorter periods of punctuation. Punctuation could be caused by several factors, either natural or manmade. The introduction of new technology, invasion, mass migration, fire or flood are some examples of punctuation. Dixon (1997:73) argues that it is only in times of punctuation that recognizable splits, of the kind well modeled by a family tree, occur.

A prediction of punctuated equilibrium is that languages in an equilibrium situation will 'converge to a common prototype'; that is, they will tend to borrow lexical items and grammatical structures until they reach approximately 50% in common (Dixon 2002:26-30). According to Dixon, in Australia any lexical or grammatical component of a language can be borrowed between languages, and there is there is no universal that basic vocabulary is borrowed less than non-basic (although this may be the case in European languages).

Although Dixon challenges the applicability of the family tree model to Australian languages, to my knowledge he does not challenge the principle of the regularity of sound change or the applicability of the comparative method in broad terms, although he claims that the amount of borrowing between Australian languages make its application susceptible to false results. The comparative method is based on the assumption that internal language change is regular, and that regular, systematic correspondences between languages are meaningful and highly unlikely to be due to chance. What one does with this inference, however, is not strictly part of the comparative method.[2] The comparative method only relies on our assumption that languages employ regular sound change and exceptions to this require special pleading. If Australian languages could be shown to change sporadically, then this would mean that the comparative method could not be applied in Australia, although nowhere in Rise and Fall of Languages

[2] Nor, incidentally, is the comparative method necessarily related to the family tree; this point will be taken up further below. The comparative method can still be used, for example, within linguistic areas and in the analysis of loan words (for an Australian example see Koch 1997 on loan words between Arandic and Warlpiri), which is what makes it such a powerful tool.

(or elsewhere, to my knowledge) does Dixon challenge the idea of regularity of sound change in principle.

Dixon (1997:149-152) does attack the comparative method as a 'discovery procedure', lambasting those (such as Hoenigswald 1960 and Anttila 1989) that equate the comparative method to an automatic synchronic phonological analysis. He argues that there are many places where a strict application of the comparative method will yield the wrong result. He is of course right to attack such reconstructions, although one would be hard put to find a thorough, detailed and plausible reconstruction of any family using the comparative method that applied it blindly, without taking into consideration the likelihood of the changes to be reconstructed and the possibility of language contact. In reality, neither synchronic nor diachronic phonology is practiced as a mechanical discovery procedure. Problems like opacity are encountered, recognized, and dealt with in both synchronic and diachronic linguistics.

Dixon further criticizes the comparative method because it will over-generate in some cases. For example, if a change occurs in all the languages of a group independently, it may well be reconstructed to the ancestor of the group, giving a false common innovation. We cannot reconstruct states for which we have no evidence. This is, of course, a weakness of the comparative method; however, it is a problem with all reconstructions of the past — if evidence has been obliterated, we won't see it, no matter what the method used (carbon-dating, stratigraphy, thermo-luminescence, punctuated equilibrium, glottochronology). This isn't a valid criticism of the comparative method *per se*, it's a fact of prehistory.

Dixon's third criticism of the comparative method is that proto-languages look neat and uniform, whereas 'natural languages are rather like an old garment that has been patched and mended' Dixon (1997:45). Dixon implies here and elsewhere that the comparative method can only reconstruct regularity and that we can never reconstruct all aspects of a proto-language. I do not dispute that we will never be able to reconstruct all aspects of a proto language, but I do not agree that we can only reconstruct regularity. To take some established examples from Indo-European of areas we where can reconstruct irregularity: we can reconstruct certain irregularities in paradigms, such as the presence of full grades in locatives of ablauting root nouns despite the expected zero grade for oblique cases. We can reconstruct a vowel *a distinct from that of the sequence *h_2e (that is, the *a*-coloring laryngeal + *e*) despite their merger in all daughter languages: compare the behavior of the Greek root δαμάζω 'I tame', and its ø-grade passive participle δμᾱτός (< *damH-/dmH-tó-), which shows a 'real' *a* which deletes in the zero grade, with the reflex of the root *steh$_2$- 'stand' which does not lose the vowel but gives στατός, from *sth$_2$-tó-, with vocalization of the laryngeal. Another example of reconstructing irregularity involves the remodeling of forms of the nominative case in some paradigms, due to Szemerenyi's Law (involving compensatory lengthening over a resonant as the result of the loss of final *-*s*). There are numerous other minor 'rules' to deal *precisely* with the irregularities which we can reconstruct.

Another 'reconstruction of irregularity' which Dixon does not mention is loan-word analysis. It is precisely because these forms are irregular that we can identify them as being worthy of separate explanation. It is also possible to reconstruct loans into proto-languages, using with case the same methods which we use to identify loans in modern languages.[3] For example, it may be the case that the loan from B to A shows a different reconstructible accent pattern, an odd consonant cluster, a morpheme otherwise unidentifiable, or an odd root shape. Such criteria allow us to infer, for example, that Proto-Germanic *rīk(i)ja- (Gothic *reikeis*, Old High German *rihhi*, English *rich*) is a loan from Celtic. Comparison with Latin *rex* and Sanskrit *rāj-* imply a reconstruction **rēg-*, and Celtic exhibits a regular sound change **ē* > *i*, but this is not a regular change in Germanic. The irregular correspondence identifies the word as a probable loan. Thus instances where the comparative method fails can be as instructive as instances where it succeeds.

3. Punctuated equilibrium in Australia

Dixon (1997:91) regards Australia as a 'prototypical example of a long-time diffusion area'. Indeed, he states that his punctuated equilibrium model is the *only* way to account for the current distribution of Australian languages. In this section I outline the main arguments that Dixon presents, specifically applied to Australian languages.[4]

Dixon's views of several aspects of Australian languages are at variance from those of most other Australianists (as I understand them) and go some way to explaining his insistence on the lack of applicability of the family tree model in Australia. The main difference is the view that non-Pama-Nyungan languages are archaic in relation to Proto-Australian, and the non-Pama-Nyungan languages are innovative. This follows in part from the fact that Dixon's 2002 reconstructions follow general typological principles — that is, that languages tend to go from isolating to agglutinative to inflectional and back again; this is the basis for his view of the non-Pama-Nyungan languages as innovative and the Proto Pama-Nyungan 'area' as an archaic diffusion area. The Pama-Nyungan type is assumed to be the original type, and the non-Pama-Nyungan languages are assumed to have innovated from that type by cliticizing pronouns or catalysts to verbs to form inflectional bundles, as has also happened sporadically in Pama-Nyungan languages such as the Southern dialects of Baagandji (Hercus 1986).

Dixon also assumes that Australian languages have been evolving fairly much *in situ* from early on since the initial expansion into Australia (probably around 60,000 years ago). His reasons are partly to do with origin myths (cf Dixon 1996) and partly because, as he states, if Proto-Australian had been spoken only a few

[3] See, for example, Sammallahti 1998 on loans into Saamic languages from Indo-European.

[4] One cannot do justice to the full picture in few pages; the reader should consult Dixon (2002) for a much more detail account.

thousand years ago, the split and expansion would have happened rather recently, and a family tree model should be applicable (Dixon 1997:92).

Finally, Dixon presents two arguments that Pama-Nyungan languages do not form a genetic family. The first is that if Pama-Nyungan were a genetic group, we should be able to reconstruct innovations between Pama-Nyungan and higher level groups (e.g. Dixon 2001:93). The paucity of such reconstructions should cause us to question the genetic status of Pama-Nyungan itself. He further argues in the appendix of Dixon 2001 that many of the forms taken to be common retentions from earlier stages of Pama-Nyungan (for example the widespread first person dual pronoun *ngali*) should better be treated as diffusions across a large area. The wide-scale diffusion and convergence that Dixon argues for in punctuated equilibrium is also his explanation for the lack of progress in reconstructing subgroups of Pama-Nyungan and for the non-bunching isoglosses which criss-cross the continent.

4. Problems with punctuated equilibrium as applied to Australia

Dixon is correct to point out the lack of reliably-established groups and subgroups in Australia, however there are several problems with the arguments he deduces to claim that traditional methods do not work in Australia.

Much previous subgrouping has been based on lexicostatistics (especially that of O'Grady, Wurm and Hale 1966). The O'Grady, Wurm and Hale classification was based on little data (a 100-item list) and was never meant to stand as a final classification (Wurm 1972:109). It dates from a time when our basic knowledge of Australian languages was still very patchy – indeed, part of its purpose was a 'stock-take' of the languages still spoken. The O'Grady, Wurm and Hale classification, in much later Australian comparative work, has been either assumed or rejected out of hand, without proper evaluation by traditional methods. It is sometimes forgotten that it was a provisional schema, with each grouping requiring confirmation or correction by traditional historical-comparative methods. However, until recently detailed data have not been readily available to allow much accurate reconstruction to take place.

Dixon argues against the validity of a reconstructed language Proto-Pama-Nyungan partly on the grounds that one can reconstruct no innovations from a higher proto-language that set it apart from other Australian groups. One must realize, however, that demonstrating genetic relatedness and demonstrating membership of a subgroup are different things; for example, one does not need to believe in and reconstruct Proto-Nostratic in order to define Proto-Indo-European! It is true, of course, that if one views Proto-Pama-Nyungan as a subgroup of 'Proto-Australian', to justify this one must provide cognates from Proto-Australian and show innovations to Proto-Pama-Nyungan, but it is not necessary to do this to show the cohesiveness of a Pama-Nyungan family of itself.

Finally, let us consider the arguments that Dixon uses to claim that Pama-Nyungan is an ancient linguistic area. Part of the argument rests on the use of origin myths and oral history (e.g. Dixon 1996) which, if true, would tie certain

Pama-Nyungan groups to particular areas for at least the last 10,000 years. It is worth pointing out in general terms some of the problems with using origin myths and other types of myths to date population expansions and other aspects of prehistory. Take as an example the following story about a volcanic eruption, told in Dyirbal to Dixon in the 1960s and quoted from Dixon 1972:29:

> It appears that beneath the veneer of fantasy some myths may provide accurate histories of events in the distant past of the people. There is, for instance, a Ngadyan myth that explains the origin of the three volcanic crater lakes Yidyam (Lake Eacham), Barany (Lake Barrine) and Ngimun (Lake Euramoo). It is said that two newly-initiated men broke a taboo and so angered the rainbow serpent, the major spirit of the area (...). As a result, 'the camping-place began to change, the earth under the camp roaring like thunder. The wind started to blow down, as if a cyclone were coming. The camping-place began to twist and crack. While this was happening there was in the sky a red cloud, of a hue never seen before. The people tried to run from side to side but were swallowed by a crack which opened in the ground ...' This is a plausible description of a volcanic eruption. After telling the myth, in 1964, the storyteller remarked that when this happened the country round the lakes was 'not jungle - just open scrub'. In 1968, a dated pollen diagram from the organic sediments of Lake Euramoo by Peter Kershaw showed, rather surprisingly, that the rain forest in that area is only about 7,600 years old. The formation of the three volcanic lakes took place at least 10,000 years ago. All this points to the story of the volcanic eruptions, and of the spread of the rain forest, having been handed down from generation to generation for something like ten millennia. This is perfectly possible: recent archaeological word suggests that aborigines have been in Australia for at least 25,000 years,[5] and the Dyirbal could well have been in more or less their present territory for 10,000 years or more.

There are several problems with the assumption that Dixon draws. First, just because people tell a story now and claim it as their own, that should not imply that they are the direct descendants of the people that the story happened to. Dixon gives no dates for population settlements in the areas he talks about, but from Mulvaney and Kamminga (1999:334) the area was settled around 5000 years ago, and extensive exploitation began from around 2000 years ago, considerably later than the eruption.

Furthermore, the events in such stories are often very general and do not need to have been true (for example, seeing an island off the coast, it is very easy to make up a story that it was once possible to walk to it). The events, even if true, might not refer to the same area. Australia is, after all, a very big place.

More importantly for methodology, for every element in a story that might be true, there is usually a chunk of the story that is discarded (for example, the

[5] The figure is now 55,000 years +.

rainbow serpent in the Dyirbal myth above). This very selective use of data leads us to ignore almost *everything* which does not fit our ideas of the prehistory of the area and only pick up on the aspects of the myth that fit.

To return briefly to archaeology, we do have evidence for settlement beginning well over 40,000 years ago in different parts of the country (O'Connor 1999, Mulvaney and Kamminga 1999). However, we also have ample evidence that the current distribution of languages is highly unlikely to represent anything like patterns of original settlement. The archeological literature is full of descriptions of site abandonment and recolonization over time periods of over 10,000 years (see, for example, Smith *et al* 1991). We also have considerable evidence for Holocene expansion, intensification and reoccupation of previously abandoned areas (McConvell 1996, Veth 1993, for example). Thus we have no evidence that the distribution of Pama-Nyungan languages reflects the dispersal of the original settlers to Australia, but we do have quite a bit of evidence for a *lack* of continuity in many parts of the country.

5. An alternative model

I have gone into a discussion of punctuated equilibrium in some detail, in order to describe the model and to summarize the relevant issues in accounting for the current distribution of Australian languages. In this section, I outline my alternative scenario and the accompanying model of linguistic differentiation as it applies to the Australian situation.

The ideas on which this model is based have been around for quite a while, and various authors, including McConvell and Evans (1997), assume something like what I am proposing here. The model of that describes linguistic differentiation is due almost entirely to Ross (1988, 1996, 1997). However, it is worth spelling out the scenario in detail, and making a number of assumptions explicit, precisely for the reason that this model does seem to be assumed or subsumed under a family tree model. Evans (in press), McConvell and others who have written on the question still write in terms of a splitting family tree which is as yet to be fully articulated.

Another important point to make is that this model is a model of language change and classification/genetic relatedness; it is not a model encompassing socio-historical factors (other than the point that speakers of different linguistic allegiances must remain in contact); thus while I subscribe to the view (following Thomason and Kaufman 1988) that the history of a group of languages is a function of the history of its speakers, and does not exist in isolation, I do believe that socio-historical factors should be modeled independently of their linguistic consequences; as Kuteva (1999) and Campbell (in press a, in press b) have convincingly shown, there is no one-to-one correspondence between non-linguistic events and types of language change; thus while linguistic history is a function of non-linguistic events, the two can, and should, in my opinion, be modeled separately.[6]

Finally, what I am suggesting here is one possible way in which the languages of Australia, particularly of the Pama-Nyungan group, could be related, and a scenario to account for their current distribution. I am not claiming by any means that it is the only way, but it does fit our current knowledge not only of the Australian prehistoric situation, but our knowledge of mechanisms of language change elsewhere in the world — in my model, Australia is an unusual case, but it does not require special pleading.[7]

The scenario that I propose to account for the lack of clear binary splits in Pama-Nyungan has its basis in dialect geography. The principles of dialect geography are well-known from works such as Chambers (1998) and will only be briefly repeated here. The three most important for my purposes are:

- Speech communities are differentiated by isoglosses; isoglosses tend to bunch along natural barriers to communication but do not necessarily do so;
- speech communities in the same geographical area often form chains of mutual intelligibility; that is, adjacent languages may differ minimally but either ends of the continuum may be mutually unintelligible;
- speech varieties at the epicenter of an area tend to be innovative, while isolated conservative pockets exist on the fringes.[8]

Now, consider the relationship between dialects and the family tree model; it has long been a paradox of genetic linguistics that the relationships between languages may be modeled on a family tree, but the relationships between languages do not fit a tree so easily. See, for example, Hock (1991:432ff) especially p 450:

> The linguistic relationship between neighboring dialects of the same language very commonly cannot be stated in terms of tree diagrams. This is a consequence of the fact that these speech varieties remain mutually intelligible, stay in close contact, and therefore continue to interact with each other on a day-to-day basis, with shifting realignments as political and social circumstances change. It is therefore unrealistic to expect clear, 'tree-diagram' splits in such dialect continua.

---

[6] Of course, this is not to say that a unified treatment is impossible, just that it is impossible on the current data; this raises an important point about the design of linguistic models and what they do; they do not have to provide reasons for the effects exhibited in the data. There is a crucial difference between descriptive models and explanatory models, and many authors appear to confuse the two, or provide one for the other. This is (almost) purely a descriptive model.

[7] Note that although, especially in the following sections, I am directly contrasting my model with Dixon's; I do not mean this to be an adversarial paper, far from it! Rather, I wish simply to point out that there are strikingly different ways of viewing the same data.

[8] Note that the 'epicenter' of a linguistic area need not be the geographical center of the area.

Now, consider what we would expect to happen if dialectal speech communities continued to diverge, *in situ* and remaining in linguistic and social contact with one another. I argue that in such a situation we would expect to find a series of dialectal epicenters, with the speech communities around those epicenters speaking varieties more similar to one another's than to varieties of other epicenters; we would expect to find some isoglossic scarps; that is, divisions where isoglosses have accumulated and bunch; we would also expect to find conflicting subgroupings. That is, we would find languages that exhibit traits that would in family tree terms lead them to be classified in multiple groups.

In this scenario, the possibilities for subgrouping would be quite limited, and often conflicting. However, crucially, the languages are still *genetically related*, in the same way that dialects without bunching isoglosses are. They will still probably exhibit regular correspondences and they are still descended from a single parent.

There seem to be three major advantages to this model over punctuated equilibrium to account for the distribution of languages in Australia. Firstly, there is a place for both divergence and convergence as processes of language change; punctuated equilibrium stresses convergence as the main mechanism of language change. Secondly, it makes Pama-Nyungan look much more similar to other areas of the world. We no longer have to assume that Australia is a special case. Thirdly, and related to this, we do not have to assume in this model that there has been intensive diffusion of many linguistic elements that in other parts of the world are resistant to borrowing (such as shared irregularities).

My model does assume that the dispersal of speakers of Pama-Nyungan is relatively recent; not, that is, of the order of 50,000 years or more ago, perhaps an order of magnitude more recent. This is in accordance with the relevant archaeological evidence, although, of course, linking the spread of Pama-Nyungan with the series of (re)colonizations of abandoned or previously uninhabited areas is an unsupported assumption.

These claims should not be considered original, and others have claimed what amounts almost to the same thing. The work of Malcolm Ross, in particular, discusses cases of conflicting subgrouping in Oceanic and in particular 'innovation-linked' subgroups; such groups share a common set of innovations, which are not shared outside the subgroup, yet the languages themselves provide us with no evidence for a single common proto-languages from which all the languages in the subgroup are descended.

6. Conclusions and further directions

I have shown in this paper that the punctuated equilibrium model and the classical family tree models are not the only way to view the history of Pama-Nyungan expansion in Australia; that we can assume a date of expansion more recent than initial colonization and still account for the lack of an articulated family tree, and that we can reconcile diffusion areas with recent migration and still talk in terms of genetic relationship.

Of course, this model requires intensive testing, both at the level of Pama-Nyungan but also at the level of subgroups of Pama-Nyungan. Work on this is progressing, including the results published in Bowern and Koch (2002) and Dench (2001); other work is in progress. Without detailed, step-by-step reconstruction and plotting of isoglosses, hypotheses of innovations, and the like, however, these ideas remain untestable.

I show that the current distribution of Australian languages need not be explained by the heavy diffusion scenario that Dixon argues for, but can be explained within mainstream models of language change. Indeed, I argue that the Pama-Nyungan family can be most easily explained as a relic of a dialect area. This implies that the similarities between Pama-Nyungan languages are due primarily to common genetic inheritance, and the overlapping isoglosses, to a large extent, from either parallel innovation or from old dialect areas.

References

Aikhenvald, Alexandra and R.M.W. Dixon. (eds.) 2001. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press.

Anttila, Raimo. 1989/1972. *Historical and Comparative Linguistics. Current Issues in Linguistic Theory*, vol 6. Amsterdam: John Benjamins.

Bowern, Claire and Harold Koch (eds.) 2002. *Subgrouping and Reconstruction: Case Studies from Australian Languages*. ms.

Campbell, Lyle. in press a. Beyond the comparative method? In Barry Blake and Kate Burridge (eds.), *Selected papers from the Fifteenth International conference on historical linguistics*, Amsterdam: John Benjamins.

Campbell, Lyle. in press b. What drives linguistic diversity and language spread? In Colin Renfrew and Peter Bellwood (eds.), *Language-Farming Dispersals*, Cambridge: McDonald Institute for Archaeological Research, p. 22 pp.

Chambers, J K. 1998. *Dialectology*. New York: Cambridge University Press, 2nd edn.

Dench, Alan. 2001. Descent and diffusion: the complexity of the Pilbara situation. In Aikhenvald and Dixon, chap. 5, pp. 105–133.

Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland. Cambridge Studies in Linguistics,* vol 9. London: Cambridge University Press.

Dixon, R.M.W. 1996. Origin legends and linguistic relationships. *Oceania* 67: 127–139.

Dixon, R.M.W. 1997. *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 2001. The Australian linguistic area. In Aikhenvald and Dixon, chap. 4, pp. 64–104.

Dixon, R.M.W. 2002. *Australian Languages: Their Nature and Development*. Cambridge: Cambridge University Press.

Eldredge, Niles and Stephen Jay Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In T J M Schopf (ed.), *Models in paleobiology*, San Francisco: Freeman, Cooper, pp. 82–115.

Evans, Nicholas (ed.) in press. *Studies in comparative Non-Pama-Nyungan*. Canberra: Pacific Linguistics.

Hercus, Luise Anna. 1986. *The Baagandji language*. Canberra: Pacific Linguistics.

Hock, Hans Henrich. 1991. *Principles of historical linguistics*. Berlin: Mouton de Gruyter, 2nd edn.

Hoenigswald, Harry M. 1960. *Language change and linguistic reconstruction*. Chicago: University of Chicago Press.

Koch, Harold. 1997. Comparative linguistics and Australian prehistory. In McConvell and Evans, chap. 3, pp. 27–43.

Kuteva, Tania. 1999. Language and societies: the "punctuated equilibrium" model of language development. *Language and Communication* 19: 213–228.

McConvell, Patrick. 1996. Backtracking to Babel: the chronology of Pama-Nyungan expansion in Austalia. *Archaeology in Oceania* 31:125–44.

McConvell, Patrick and Nicholas Evans. 1997. *Archaeology and linguistics: Aboriginal Australia in global perspective*. Melbourne: Melbourne University Press.

Mulvaney, John and J. Kamminga. 1999. *The prehistory of Australia*. London: Smithsonian Institution Press.

O'Grady, Geoffrey, Kenneth Hale and Stephen Wurm. 1966. Languages of Australia. Map, University of Victoria, BC.

O'Connor, Sue. 1999. *30,000 years of Aboriginal occupation*. Terra Australis: vol 14.

Ross, Malcolm D. 1988. *Proto-Oceanic and the Austronesian languages of Western Melanesia*. C98. Canberra: Pacific Linguistics.

Ross, Malcolm. 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In Mark Durie and Malcolm Ross (eds.) *The comparative method reviewed : regularity and irregularity in language change*. New York: Oxford University Press. Chapter. 8, pp. 180–217.

Ross, Malcolm D. 1997. Social networks and kinds of speech-community event. In Roger Blench and Matthew Spriggs (eds.), *Archaeology and Language I*, London: Routledge, chap. 13, pp. 209–261.

Sammallahti, Pekka (1998). The Saami languages: an introduction. Kárášjohka: Davvi Girji.

Smith, Moya, E. Williams and R.J Wasson 1991. The archaeology of the JSN site: some implications for the synamics of human occupation in the Strzlecki desert during the late Pleistocene. *Records of the South Australian Museum* 25(II):175-192.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley and Los Angeles: University of California Press.

Veth, Peter. 1993. *Islands in the interior.* International Monographs in Prehistory. Archaeological Series 3, University of Michigan, Ann Arbor.
Wurm, Stephen A. 1972. *Languages of Australia and Tasmania.* the Hague: Mouton de Gruyter.

Department of Linguistics
Harvard University
305 Boylston Hall
Cambridge, MA, 02138
USA

bowern@fas.harvard.edu

Quantity, causality and temporality in change constructions

CRISTIANO BROCCIAS
Università di Genova

0. Introduction

This paper investigates within a Cognitive Grammar framework three English constructions which are regarded as instantiations of a more general construction, the change construction. They evoke the notions of quantity, causality, and temporality, respectively, and it is apparent from the use of dynamic spatial prepositions such as *to* and *into* that they are construed as paths. The three English constructions are also compared with their Italian equivalents, thus allowing us to draw some general conclusions about the availability of change constructions in the two languages.

1. Three English constructions involving path construal

1.1. Quantification as a path

Quantification can be conceptualised as a path, as is shown by the bracketed prepositional phrases in (1)-(3) below.

(1) a. I love you [all the way up **to** my toes].
b. I love you [all the way **down** the lane as **far** as the river].
c. I love you [**across** the river and **over** the hills].
d. I love you [right up **to** the moon and back].

(2) How do I love thee? Let me count the ways.
I love thee [**to** the depth and breadth and height
My soul can reach], ...
I love thee [**to** the level of everyday's
most quiet need], by sun and candlelight.
(Elizabeth Barrett Browning, *Sonnets from the Portuguese*, Sonnet 43, ll.1-3 and 5-6)

(3) I love you [**to** heaven/**to** the power of everything plus one]. (Stephen Fry, *The Stars' Tennis Balls*, p.11 and p.19)

Let us consider the examples in (1), from a children's book by Sam McBratney (*Guess How Much I Love You*), as a matter of illustration. The intensity of the subject referent's love for the object referent is visualised as a path from the speaker (and hearer)'s location to some progressively farther and farther away landmark (or target). Similar considerations hold good for (2) and (3), which therefore illustrate that the pattern under scrutiny is, at least to some extent, productive in English.

1.2. Causality as a path

The examples in (4) demonstrate that causality can also be conceptualised as a path. For instance, (4a) implies a causal relation between the verbal event and the coming about of the librarians' astonishment. Crucially, such a relation undergoes path construal as is signalled by the spatial preposition *to*. In other words, the verbal event *led to* the psychological state denoted by the object of the preposition.[1]

(4) a. [S]he brazenly entered her name in the ledger as Erasmus Darwin ... [**to** the astonishment of the *librarians*]. (Alberto Manguel, *A History of Reading*, p.318)

b. It was the Christmas season ... and Winifred got trapped in the carol singing, [much **to** *her* annoyance]. (Margaret Atwood, *The Blind Assassin*, pp.533-534)

c. After that, they can kill everyone [**to** *their heart*'s content], if that's what they want to do. (Margaret Atwood, *The Blind Assassin*, p.421)

d. Clothes could always be purchased, naturally, but I would have to learn to wear them [**to** effect]. (Margaret Atwood, *The Blind Assassin*, p.285)

e. Then Mrs Coulter spoke, [**to** *their* surprise]. (Philip Pullman, *The Amber Spyglass*, p.214)

f. Eventually the hammer was formed [**to** *Iorek*'s satisfaction]. (Philip Pullman, *The Amber Spyglass*, p.198)

g. "Yes, that is what I heard," said the priest, [**to** *Will*'s relief]. (Philip Pullman, *The Amber Spyglass*, p.105)

1.3. Temporality as a path

The last construction to be analysed here is exemplified in (5a), which is to be contrasted with (5b). Instead of the temporal preposition *until*, the prototypically dynamic spatial preposition *into* can be used to code temporality:

[1] Observe that this construction usually has a psychological deverbal noun (underlined in (1)) as the object of the preposition and that the psychologically affected entity (italicised in (1)) is usually expressed (but see (1d), where it is not since it has a non-specific reading).

(5)	a.	She read well **into** the night.
	b.	She read **until** late at night.

The three constructions exemplified above are also very interesting from a cross-linguistic point of view. As a matter of illustration, let us consider the translations of (1d), (2), a simplified version of (4a), and (5a):

(6)	Ti	voglio	bene	{??**fino** alla/***a**lla}	luna,	lassù,	e	ritorno.
	(to)-you	(I)-want	good	until at-the/at-the	moon	there-up	and	return

(7)	Ti	amo	**fino**	**a**gli	estremi	di	profondità, ...		
	you	(I)-love	until	at-the	limits	of	depth		
	Ti	amo	**entro**	la	sfera	delle	necessità	quotidiane...	
	you	(I)-love	within	the	sphere	of-the	needs	everyday (adj.)	

(8)	Si	firmò	come	E. D.	{**per** lo/**con**/***a**llo}	stupore
	self	(s/he)-signed	as	E. D.	{**for** the/**with**/***at**-the}	astonishment
	dei	bibliotecari				
	of-the	librarians				

(9)	a.	* Lesse	(ben)	**nella**	notte.	
		(s/he)-read	(well)	in-the	night	
	b.	Lesse	**fino**	**a**	notte	fonda.
		(s/he)-read	until	at	night	deep (i.e. the dead of night)

The Italian data show that prototypically simplex spatial prepositions like *a* (cf. English *at*) and *in* (which is realised as *nella* if followed by the determiner *la*, as in (9a)) cannot be used. Further, a complex preposition like *fino a*[2] leads to variation in native speakers' judgements. (6) is very awkward (despite appearing in the Italian translation of McBratney's book). The occurrence of *fino a* in (7) sounds fine but may be influenced by both a non-spatial reading (cf. *Ti amo fino alla morte*, "I love you to death", lit. "until at the death") and the length of the prepositional phrase. Interestingly, the second occurrence of the original *to* (i.e. *to the depth*) has been replaced with an altogether different preposition, namely *entro* (cf. English *within*), which evokes a container-contained relationship. Similarly, *a* is impossible in (8), where English *to* can be translated as either *per* (cf. English *for*) or *con* (cf. English *with*). Finally, *fino a* is perfect in (9b) but because it retains its usual temporal meaning.

The three English constructions under scrutiny and their often awkward or impossible word-by-word renderings into Italian pose two interrelated questions,

[2] I have glossed *fino a* as *until at* rather than *up to* to explicitly point out the fact that *a* is not an intrinsically dynamic preposition like English *to*. That is, *a* is neutral to the distinction between motion (cf. English *to*) and lack thereof (cf. English *at*). Historically, however, *a* derives from the dynamic Latin preposition *a(d)*, which takes accusative case.

which I will address below. First, are the quantification as a path construction (QPC), the causality as a path construction (CPC), and the temporality as a path construction (TPC) related to other English constructions which similarly employ spatial conceptualisation and turn out to be (usually) impossible in Italian? Second, given that prepositions are relational predications and hence involve a trajector and a landmark (using Langacker's Cognitive Grammar terminology), what element is the trajector of a preposition like *to* in the examples above?

2. Change constructions

Starting with the former question, I propose that the QPC, CPC, and TPC are systemically related to the so-called resultative construction (see Goldberg 1995, Rappaport Hovav and Levin 2001 among many others), which I take here as also subsuming Goldberg's Caused Motion Construction (i.e. I do not distinguish in principle between states and positions, see also Broccias 2001). In more detail, both the QPC, CPC, and TPC, on the one hand, and the resultative construction, on the other, can be said to originate from the blending (in the sense of Fauconnier and Turner 2002) of two input spaces (or components), an event space and a change (or path) space. The former is coded by the verb and describes the involvement of an entity in an event E; the latter is coded by the *change phrase* (CP), which is a non-verbal phrase (corresponding to a prepositional phrase in the cases at hand), and depicts a change of state or position. The resulting clause is called a change construction (CC for short; see Broccias *forthcoming*). Consider (10) and (11):

(10) He drank himself [to death]$_{CP}$. (*blend*)
Input 1 (event space): drinking event
Input 2 (change space): dying event

(11) I love you [right up to the moon]$_{CP}$. (*blend*)
Input 1: loving event
Input 2: motion (to the moon) event

Both (10), a resultative construction, and (11), a QPC, can be analysed as stemming from the blending of two input spaces. (10) merges the drinking event with the event of somebody's dying (as a result of it). (11) merges the event of somebody's loving somebody else with a motion-to-the-moon event. The crucial difference between (10) and (11) involves the notion of causality. Whereas drinking is conducive to the subject referent's death in (10), no causal relation exists between the two constitutive components of (11). In other words, change constructions can be either causal or noncausal. Change constructions can be further subdivided into lexical and sublexical ones, as is shown in (12):

(12) a. She punched him$_i$ [to death]$_i$. (causal lexical CC)
b. He fired the gun [into the supermarket]. (causal sublexical CC)

c. She$_i$ walked [into the room]$_i$. (noncausal lexical CC)

The CP *to death* is predicated of an overtly realised element in (12a) as indicated by the subscript index *i*. This is not the case in (12b), where the CP is understood as referring to the bullets "emitted" by the gun. Since the emitted substance (i.e. the bullets) are not specified in the syntax, I will refer to it as a sublexical entity (see Broccias 2001) and say that the construction is sublexically oriented. Finally, (12c) illustrates a noncausal lexical CC in that *into the room* is (usually) interpreted as simply specifying the location arrived at by the subject referent (see Broccias *forthcoming*: chapter 2 for details on the causative reading).

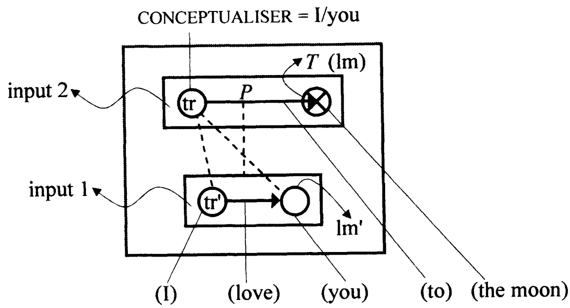
The paradigm in (12) obviously includes a gap since it does not contain any noncausal sublexical CC and hence poses the question of whether such a construction exists at all. My claim, which I will defend below, is that (a) the QPC and the CPC are indeed instantiations of the noncausal sublexical CC and (b) the TPC is a realisation of the noncausal lexical CC.

3. The quantification as a path construction as a change construction

It is worth observing that analogy may be insufficient to motivate the existence of the QPC. One could argue that a sentence like *I love you to the moon and back* has been modelled onto well-entrenched examples like *I love you to {distraction/tears/bits and pieces}*. The relevant prepositional phrases (which evoke an above-the-norm reading) may simply have been substituted with a *spatial* prepositional phrase. Still, if this were the end of the story, one would expect its Italian equivalent to be possible, given the existence of examples similar to the English ones such as *Ti amo alla follia* (lit. "I love you at the folly", i.e. "I love you to distraction"). But, as was shown above (see (6a)), the use of Italian *at* is banned and that of *fino a* is problematic. It may well be the case that the greater productivity of the QPC in English as compared to Italian rests on the greater productivity of the (more general) blending mechanism illustrated in section 2 for change constructions. In other words, analogy may need to be supplemented by a readily available grammatical construction, that is a high activation potential for the change construction (intended as the grammatical output of the blending operation of section 2). This in turn may be related to the greater availability of dynamic prepositions in English (see section 6).

Although the dynamic preposition *to* in *I love you to the moon and back* explicitly signals a path (i.e. the prepositional phrase it heads counts as a change phrase), no causal relation exists between the verbal event and the change event (see section 2). Hence, the relevant change construction is to be classified as noncausal. It must now be decided what entity the change phrase is predicated of. Apparently, the change phrase is predicated of the subject and object referents: the speaker (*I*) invites the addressee (*you*) to scan along and thus measure a path so that the latter might compare it with the extension of love as a reified object. In other words, the non-spatial relation denoted by *love* is projected onto a path scanned by the speaker and addressee arriving at the location specified as the preposition's complement. This has been visualised in (13) for a simplified version of the sentence under discussion.

(13)



(13) shows that *I love you to the moon* can be analysed as the blending of two input spaces (i.e. input 1 and input 2). Input 1 depicts an interaction between a trajector (*tr'*), *I*, and a landmark (*lm'*), *you*, which is symbolised as *love*. Input 2 is equated with a path scenario, where an entity (*tr*) moves along the path *P* (symbolised by *to*) thus ending up in the target (or landmark) *T* (i.e. the moon). Further, the parallel arrangement of the two components is intended to capture the lack of causality between them. The diagram explicitly indicates that the loving event (i.e. the bottom box) is projected onto (i.e. is conceptualised as) a path: a dashed correspondence line connects it to *P*. Finally, *tr* is equated with both *tr'* and *lm'* thus accounting for the interpretation that the change phrase is predicated of either.

I would like to contend that the lexical orientation for *I love you to the moon and back* is only an accident, due to the use of first and second person pronouns. If we substitute third person pronouns for them, as in *He loves her to the moon and back*, we are bound to conclude that the change phrase may not be predicated of either the subject or the object referent (or both). Rather, the conceptualisation of love as a path is due to the speaker or, in Cognitive Grammar's terminology, conceptualiser. Hence, we have an instance of what Talmy (1996) calls fictive (rather than factive) motion and Langacker (1990) subjective motion. The change phrase is predicated of the conceptualiser, who traces a mental path (see also Lee 2001: 45 for some related examples). I conclude that the change construction is noncausal and sublexical (in particular *conceptualiser-oriented*). Since the conceptualiser can be brought "on stage" by using non-third person pronouns, the sublexical nature of the construction may be blurred at first sight.

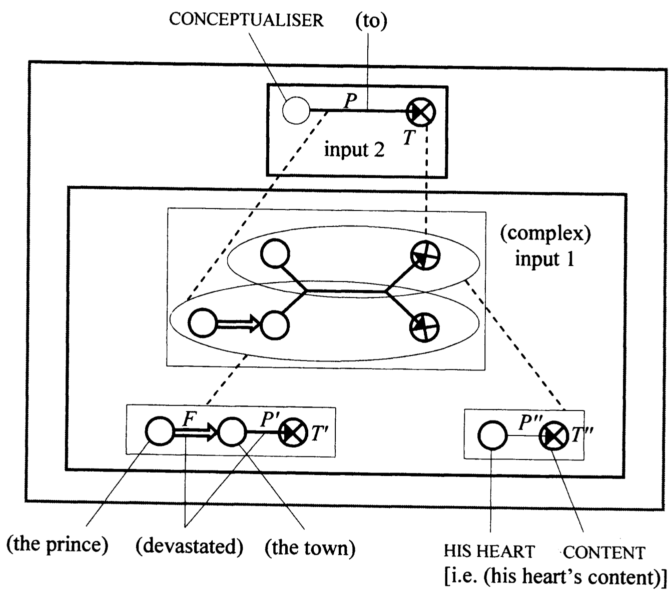
4. The causality as a path construction as a change construction

Let us now move to a more detailed analysis of the causality as a path construction, of which I will take (14) below as a representative example:

(14) The prince devastated the town [**to** his heart's content]$_{CP}$.

It was observed in section 2 that, differently from what is the case for the QPC, causality obtains between the verbal event (i.e. the prince's destruction of the town) and the complement of the preposition *to* in the CPC. Still, the trajector of the preposition *to* is not, as in "ordinary" resultatives like *She punched him to death* (see (12)), an argument of the verb. In (12), the referent of *him* moved, metaphorically speaking, *to death*. In (14), we cannot say that *the prince* moved *to his heart's content* (this is even clearer if we change the reference of the possessive determiner, as in *The prince devastated the town to **my** heart's content*). Further, the sentence is perfect in the passive (i.e. *The town was destroyed to his heart's content*) contrary to what happens with resultative cases (cf. *She$_i$ danced the waltz [into the room]$_i$* vs. **The waltz was danced into the room*; see Rappaport Hovav and Levin 2001 for some discussion). I conclude that, although the blending of the event of the prince's devastating the town and that of his heart's becoming content is a causal one, such a merger is not symbolised syntactically (via the preposition *to*). Rather, the preposition *to* symbolises a path *P* onto which (subparts of the) events of the prince's devastating the town and his heart's becoming content are projected, as is visualised in (15).

(15)



(15) shows that (14) can be treated in similar fashion to *I love you to the moon and back*, that is, it can be said to originate from the noncausal blending of two input spaces (i.e. input 1 and input 2). They differ, however, in that input 1 in (15) is complex, being made up of two subcomponents (cf. the two boxes below). The

subcomponent in the left-hand corner depicts an energy flow (*F*) from *the prince* to *the town* resulting in the change of state of the latter (the final state is *T'* in the diagram). The causative verb *devastate* of course symbolises both force exertion and change (i.e. *P'*). The subcomponent in the right-hand corner describes the change of state of the prince's heart, which reaches the state diagrammed as *T''* (i.e. the state of being content). Observe that path *P''*, standing for such a change-of-state event, has not been emboldened because it is not symbolised syntactically (contrary to what would be the case in an "ordinary" resultative construction, see Broccias 2001 for details). Further, *his heart* and *content* have been capitalised (i.e. they have been treated as conceptual entities only) in order to show that they are symbolised not at the level of input 1 but only after the merger of input 1 with input 2. Finally, the existence of a causal relation between the two subcomponents in input 1 has been diagrammed by drawing them in sequential rather than parallel fashion.

The diagram in the middle in (15) visualises the fusion of the lower two subcomponents as involving the projection of *P'* and *P''* onto each other. This is intended to capture the intuition that the event of the town's destruction and that of the prince's heart's change of state were coextensive to some degree. Crucially, the devastation event is projected onto path *P* in input 2 (which is analogous to input 2 in (13)). That is, the event of the prince's devastating the town can be viewed as a path leading to an abstract region (*T*) which corresponds to *T''* (i.e. the final state achieved by the prince's heart). As was the case in the diagram in (13), the conceptualiser can be taken as the entity mentally tracing *P* and hence corresponds to the trajector of the preposition *to*.

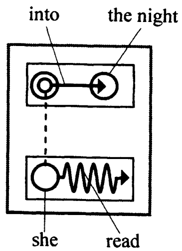
5. The temporality as a path construction as a change construction

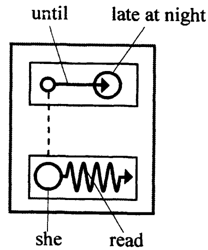
The last construction to be detailed here is the TPC, which is exemplified in (16) and illustrated in (17a) for (16a). (17a) should be considered alongside (17b), which represents the semantic import of *She read until late at night* (see (5b)).

(16) a. She read well into the night. (=(5a))
b. The chauffeur was relaxing **into** his monologue. (Ian McEwan, *The Child in Time*, p.140)

(17a) shows that the TPC is also a noncausal change construction (as indicated by the parallel arrangement of its two input spaces). Still, the TPC differs from the QPC and the CPC in that the blending of its two inputs also relies on the correspondence established between the trajector within input 1 (i.e. *she*) and the trajector within input 2 (the bigger circle in which it is inscribed depicts a source location *S*, standing for the location/state occupied by the trajector when the event designated by the verbal event starts). To put it differently, the TPC is not a conceptualiser oriented construction since the trajector of *into* is obligatorily (and not optionally as in the QPC) an argument of the verb.

(17) a. b.





(17a) visualises our (metaphorical) conceptualisation of animate beings as entities moving through time (cf. *I'm halfway through this book*, where the spatial preposition *through* is employed to code temporality). I contend that this strategy is similar to the one we resort to when we use the *until* variant, see (17b). The *until* variant is analysed here as involving the correspondence of the trajector in input 2 with the whole input 1 (as indicated by the dashed line in (17b)) rather than a subpart of it, as in (17a). In other words, an event rather than an animate being is construed as moving along the time arrow.

6. The limited availability of change constructions in Italian

The virtual lack of the QPC, CPC, and TPC in Italian has been linked (see section 3) to the limited availability of change constructions in such a language. A motive must now be found for why Italian generally lacks change constructions. It is worth observing, first of all, that Italian does have some Resultative Constructions in Goldberg's (1995) sense (i.e. resultative constructions having to do with states rather than positions) with the simplex preposition *a* (see also Merlo 1989 and Napoli 1992 for some discussion). However, their productivity is severely restricted to a few cases like (18) below:

(18) a. Ti amo **alla** follia.
you (I)-love at-the folly
'I love you to distraction'

b. Lo hanno picchiato **a** morte/sangue.
him (they)-have beaten at death/blood

As for Caused Motion Constructions (i.e. resultative constructions having to do with positions rather than states), it seems that simple prepositions like *a* and *in* (which, like *a*, is not inherently dynamic and hence may correspond to either English *in* or English *into*) cannot (usually) be used with non-punctual manner-of-motion verbs. Rather, the complex prepositions *fino a* and *fin dentro* are used. Consider the following data:

(19) a. Camminò {**fino al/*al**} bosco.
(s/he)-walked {until at-the/*at-the} wood

b. **{Ha/*E'}** camminato **fino al** bosco.
{(s/he)-has/*is} walked until at-the wood

(20) a. I soldati **sono** corsi **alle** mura.
the soldiers are run at-the walls
b. Corsi **alla** stazione (in taxi).
(I)-ran at-the station (in taxi)
'I rushed to the railway station (by taxi)'

(21) a. Il gatto è saltato **nella** scatola.
the cat is jumped in-the box
b. Ha zoppicato {**fin dentro/*in**} casa. (allative reading)
(s/he)-has limped {until inside/*in} house

(19a) shows that the non-punctual manner-of-motion verb *camminare* (English *run*) can be employed in an allative construction only if the complex preposition *fino a* is selected. Interestingly, if auxiliary selection is anything to go by – i.e. selection of auxiliary *avere* (cf. English *have*) correlates with temporality, cf. *{**Ho/*sono**} letto fino a tarda notte*, lit. '{(I)-have/*(I)-am} read until at late night' – we conclude that *fino a* has both a temporal "feature" and a spatial "feature" (i.e. the allative meaning conveyed by the construction).

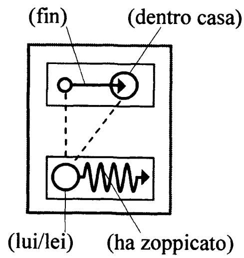
(20a) demonstrates that, in some cases, we do find manner-of-motion verbs used telically in conjunction with the simple preposition *a*. Auxiliary selection (i.e. the verb *essere*, cf. English *be*) reveals that such an example lacks temporal nuances (vs. (19a)). Nevertheless, (20a) may be a relic from Latin (cf. the infinitive form *ad muros currere*, lit. 'to walls run', quoted in Castigliotti and Mariotti 1966). Crucially, if we substitute *stazione* ('railway station') for *mura* as in (20b), the interpretation we get is that the subject referent went to the station quickly: (s)he may not have run at all but may have taken a taxi.

Finally, (21) illustrates that the simplex preposition *in* can be used only if the verb is to be interpreted punctually, as in (21a). The non-punctual verb *zoppicare* (cf. English *limp*) requires the complex preposition *fin dentro*.[3] As was the case in (19), the use of *fino* also evokes a temporal interpretation. A simplified schematic representation for (21b) is offered in (22) (*lui=he*, *lei=she*).

The diagram in (22), which must be considered alongside the two in (17), captures the intuition concerning the spatial and temporal interpretation of *fino* in (21b). The whole input 1 (the lower box) is put in correspondence with the trajector in input 2 (the upper box), as was the case for English *until*, see (17b). That is, the event of limping lasted until the target (i.e. the interior of the house) was reached. Further, the trajector within input 1 has been linked to such a target in order to show that the subject referent ended up in such a location (i.e. the allative interpretation of the construction).

[3] Final *-o* in *fino* is optional before the voiced plosive /d/.

(22)



I would like to propose that the nature of Italian simplex and complex prepositions motivates the limited productivity of change constructions in Italian. First, simplex prepositions like *a* and *in*, unlike English *to* and *into*, do not explicitly code a change (of state/position). Second, the complex preposition *fino* + simple preposition possibly evokes a complex scenario where temporality and spatiality intermingle (which may explain the non-perfect status of (6) above). These two factors may have prevented and may still be preventing the growth of Italian change constructions.

7. Conclusion

In this paper I have proposed that the availability of the English quantification as a path construction, the causality as a path construction, and the temporality as a path construction may be related to the postulation of a more general construction, the change construction, which in turn may depend on the existence of explicitly dynamic simple prepositions (contrary to what is the case in Italian). In more detail, the quantification as a path construction and the causality as a path construction have been analysed as instances of the sublexical noncausal change construction, where the preposition's trajector is to be equated with the conceptualiser. The two constructions differ however in that the latter includes a complex, rather than a simple, input 1 space. Finally, the temporality as a path construction has been regarded as an instance of the lexical noncausal change construction since the preposition's trajector necessarily (rather than optionally as in the quantification as a path construction) corresponds to the subject referent.

References

Atwood, Margaret. 2001. *The Blind Assassin*. London: Virago Press [2000].

Broccias, Cristiano. 2001. The need for the resultative network. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 26: 41-52.

Broccias, Cristiano. forthcoming. *The English Change Network*. Berlin/New York: Mouton de Gruyter.

Castigliotti, Luigi and Scevola Mariotti. 1966. *Vocabolario della lingua latina.* Torino: Loescher.

Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think. Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
Fry, Stephen. 2001. *The Stars' Tennis Balls*. London: Arrow [2000].
Goldberg, Adele. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Langacker, Ronald. 1990. *Concept, Image, and Symbol. The Cognitive Basis of Grammar*. Berlin/New York: Mouton de Gruyter.
Lee, David. 2001. *Cognitive Linguistics. An Introduction*. Oxford: Oxford University Press.
Manguel, Alberto. 1997. *A History of Reading*. London: Flamingo [1996].
McBratney, Sam. 2000. *Guess How Much I Love You*. London: Walker Books [1994].
McEwan, Ian. 1992. *The Child in Time*. London: Vintage [1987].
Merlo, Paola. 1989. Risultative in italiano e inglese. *Rivista di Grammatica Generativa* 14: 29-53.
Napoli, Donna. 1992 Secondary resultative predicates in Italian. *Journal of Linguistics* 28: 53-90.
Pullman, Philip. 2001. *The Amber Spyglass*. London: Scholastic [2000].
Rappaport Hovav, Malka and Beth Levin. 2001. An event structure account of English resultatives. *Language* 77: 766-797.
Talmy, Leonard. 1996. Fictive motion in language and "ception". In P. Bloom, M. Peterson, L. Nadel & M. Garrett (eds.). *Language and Space*, 211-276. Cambridge, MA: MIT Press.

Cristiano Broccias
Dipartimento di Scienze della
Comunicazione Linguistica e Culturale
Università di Genova
P.zza S. Sabina, 2
16125 Genova
Italy

cribroc@tin.it

Preemptive usage of N_1 *to yuu* N_2 in Japanese discourse: Psychological strategy to increase the distance between the hearer and the entity

MICHIKO TODOKORO BUCHANAN
University of Minnesota

0. Introduction

This paper examines the usage of the Japanese NP, N_1 *to yuu* N_2 'N_2 called N_1'. I introduce the notion of "mutual recognition" of the entity between the speaker and the hearer(s). I define mutual recognition as one's past experience where he/she shared the process of recognizing the entity with the hearer. This notion concerns whether the speaker and the hearer shared the sense of mutual recognition about the entity prior to the conversation rather than the mutual knowledge about the entity *per se*. This study focuses on what the speaker is aiming to do by using N_1 *to yuu* N_2 in the interaction, as well as the mechanism used to accomplish the aim.

The NP, N_1 *to yuu* N_2, is conventionally used when the speaker introduces a new proper noun when he/she does not know if the hearer can recognize it. I refer to this usage as "conventional usage". In addition to the conventional usage, we also encounter the situation where the speaker uses N_1 *to yuu* N_2 knowing that the hearer can recognize the entity, which I refer to as "preemptive usage". Previous studies on the theories relating to N_1 *to yuu* N_2 (Takubo 1989, 1992; Kamio 1990, 1994; Suzuki 1996, Takubo and Kinsui 1997, Suzuki 1998) do not discuss the difference between the conventional usage and the preemptive usage, or what the speaker's goal is in using N_1 *to yuu* N_2 in the interaction.

In the present study, I describe the structural variation of N_1 *to yuu* N_2 and review the theories relating to N_1 *to yuu* N_2. Then I examine the conventional usage, and argue that N_1 *to yuu* N_2 conventionally implicates the "lack of mutual recognition" between the speaker and the hearer(s). Based on this argument, I further investigate the use of N_1 *to yuu* N_2 in the situation where the hearer can recognize the entity, and claim that the use of N_1 *to yuu* N_2 in that situation is the speaker's strategy to preempt conflict; therefore, the term preemptive usage. I argue that the speaker accomplishes the goal of preemptive usage by increasing the psychological distance between the hearer and the entity providing the sense of the lack of mutual recognition, which N_1 *to yuu* N_2 conventionally implicates. I demonstrate that the notion of mutual recognition is pertinent to explain the mechanism of the preemptive usage of N_1 *to yuu* N_2.

1. Structural variations of N_1 *to yuu* N_2

The NP, N_1 *to yuu* N_2, whose structure is [$Noun_1$ + Quotative Particle *to* + Verb *yuu* 'to say' + $Noun_2$], has some structural variations. The quotative particle *to* has its colloquial form, which is *tte* (Miura 1974:23-24). Thus, the propositions of (1) and (2) are the same.

(1) *Watasi wa kinoo* ***Tanaka-san to yuu hito*** *ni atta.*
I TOP[1] yesterday Tanaka-Mr. QP say person DAT meet-PERFECT
'I met **a person called Mr. Tanaka** yesterday.'

(2) *Watasi wa kinoo* ***Tanaka-san tte yuu hito*** *ni atta.*
I TOP yesterday Tanaka-Mr. QP say person DAT meet-PERFECT
'I met **a person called Mr. Tanaka** yesterday.'

An N_2 can be replaced with a pronoun *no* 'the one' as in (3). Jorden (1988:171) refers to the pronoun of N_1 *to yuu no*, as "anticipatory *no*", because the hearer can reconstruct the possible original NP, N_1 *to yuu* N_2, based on the classification of N_1. In (3), the hearer can know *Tanaka-san to/tte yuu hito* 'a person called Mr. Tanaka' is the original NP of *Tanaka-san to/tte yuu no*, because *Tanaka-san* is classified as a person. Thus, the propositions of (3) and (4) are the same.

(3) ***Tanaka-san to/tte yuu no*** *wa dare desu ka.*
Tanaka-Mr. QP say one TOP who CP Q
'Who is **the one (the person) called Mr. Tanaka**?'

(4) ***Tanaka-san to/tte yuu hito*** *wa dare desu ka.*
Tanaka-Mr. QP say person TOP who CP Q
'Who is **the person called Mr. Tanaka**?'

When N_1 *to yuu* N_2 is used as a topic followed by the Topic Marker *wa* as in (3) and (4), N_1 *to yuu* N_2 *wa* can be replaced by N_1 *tte* (Miura 1974:24) as in (5). In (3) and (4), 'the person called Mr. Tanaka' is the topic; therefore, *Tanaka-san to/tte yuu hito wa* 'a person called Tanaka-TOP' can be replaced by *Tanaka-san tte*. Thus, the propositions of (3), (4), and (5) are the same.

---

[1] The following abbreviations are used in the glosses.
ACC: accusative, CP: copula, DAT: dative, EXT: extended predicate, FP: final particle, GEN: genitive, NEG: negative, PASS: passive, Q: question marker, QP: quotative particle, SUB: subject, TOP: topic marker

(5) ***Tanaka-san tte*** *dare desu ka.*
Tanaka-Mr. QP who CP Q
'Who is **the person called Mr. Tanaka**?'

2. Notions relating to N_1 *to yuu* N_2 in previous studies

2.1. Takubo's "Discourse Management"

Takubo (1989, 1992) and Takubo and Kinsui (1997) demonstrate that a linguistic form that the speaker uses to introduce a proper noun into a conversation is one of the examples reflecting how the speaker takes the hearer's knowledge into account. They refer to the process as "discourse management". They describe their theory of discourse management in terms of Fauconnier's (1985) mental space. According to Takubo and Kinsui (1997:748-749), there are two mental domains: D-domain, which is "linked to the long term memory and directly accessible", and I-Domain, which is "linked to the temporary memory and indirectly accessible". The information the speaker/hearer obtained prior to the conversation can be directly accessible, that is, the information is set up in the D-domain, and the information that is introduced and processed during the conversation is set up in the I-domain. They claim that the speaker selects a linguistic form according to which domain the information belongs taking not only his/her domains into account, but also the hearer's.

Takubo and Kinsui (1997:745) argue that even if the speaker and the hearer have mutual knowledge about the entity, the speaker cannot always use bare proper nouns. When "asymmetry in the state of mutual knowledge" exists, the speaker uses "quote marker *to-iu* 'called'" to reflect the asymmetry, which he refers to as the "metareferential use of a noun phrase" (1997:743).[2] Takubo and Kinsui's (1997:749) "Constraint on Information Transfer", which states that "[i]nformation in the I-domain cannot be transferred to the D-domain during the ongoing discourse session", explains the reason why B in (6) cannot use a bare proper noun, while A changes from the metareferential form to the bare proper noun. Since the speaker B did not have information about Mr. Tanaka, which put the entity in the I-domain, B cannot change from the metareferential form to the bare proper noun during the conversation because of the constraint. In contrast, the speaker A can change from the metareferential form to the bare proper noun, because A already had information about Mr. Tanaka prior to the conversation, which put the entity into the D-domain.

(6) A: *Boku no yuuzin ni* ***Tanaka to yuu yatu*** *ga imasu.-----*
I GEN friends among Tanaka QP say person SUB exist
'Among my friends, there is a person called Tanaka.'
B: *----Zya, sono* ***Tanaka to yuu hito*** *ni tanonde kudasai.*
then the Tanaka QP say person DAT ask-REQUEST
'Then please ask the person called Tanaka.'

---

[2] *Iu* 'to say' is an orthographical form of *yuu*. The word is often colloquially pronounced as *yuu*.

A: *Zya,* ***Tanaka-kun*** *ni* *tanomimasu.*
Then Tanaka-Mr. DAT ask
'Then [I] will ask Mr. Tanaka.'
(Takubo 1989:219, modified)

2.2. Kamio's "Territory of Information"

Kamio (1990) applies his theory of "territory of information" to some NPs. The theory is based on the "notion of psychological distance between a given piece of information and the speaker/hearer" (Kamio 1994:68), and mainly concerns the relationship between linguistic forms and the psychological distance between the information and the speaker/hearer. He claims that in Japanese there are NPs that imply that the information expressed by the NP is not in the speaker and/or the hearer's territory. He claims that N_1 *to yuu* N_2 is one of them. According to Kamio, the speaker can use N_1 *to yuu* N_2 when the information is outside the speaker's territory regardless whether the information is inside or outside the hearer's territory; however, when the information is in the speaker's territory, the speaker can use N_1 *to yuu* N_2 if the information is outside the hearer's territory.

2.3. Suzuki's "Incorporation of Information"

Based on the analysis of *tte* and *to yuu no wa* (Suzuki 1996), and *tte* and *nante* 'the likes of' (Suzuki 1998), Suzuki claims that the use of these linguistic forms reflects the distance between the speaker and the entity. Her focus is on the relationship between the linguistic forms and the entity that the speaker uses as a topic. She proposes the notion of "incorporation of information", which refers to the "process by which the speaker acquires a piece of information and internalizes the information" (Suzuki 1998:431). Suzuki's notion is built on Kamio's theory of territory of information, and she claims that "the degree of incorporation is high when the speaker has digested and integrated information into his/her belief system" and "the degree is lower when the process of incorporation is blocked for some reason" (Suzuki 1998:431). In the analysis of N *tte*, Suzuki notes that the "speaker's detachment" (19998:440) is present in all cases of N *tte.* She argues that "detachment" is a psychological distance from the entity rather than recognizability of the entity. In (7), although the speaker can recognize *kimi* 'you', the speaker uses *tte* because the information about *kimi* belongs to the hearer's territory.

(7) *Fushigina* *onna* *da* *naa,* ***kimi tte*** ***onna*** *wa.*
Strange woman CP FP you QP woman TOP
'you are a strange woman, you (lit. **a woman called you**).'
(Suzuki 1998:443, cited from Martin 1975:1017, modified)

3. Conventional usage of N_1 *to yuu* N_2 and notion of mutual recognition

I introduce the notion of mutual recognition to describe the conventional usage of N_1 *to yuu* N_2. By mutual recognition, I mean the speaker's past experience of sharing the process of recognizing the entity with the hearer. For entities that are commonly known, the mutual recognition of the entity is considered to be tacitly established prior to the conversation. I claim that when the speaker believes that he/she and the hearer did not share the process of recognizing the entity prior to the conversation, which I refer to as "lack of mutual recognition", the speaker uses N_1 *to yuu* N_2 to introduce the proper noun. Thus, for example, when the speaker does not believe that prior to the conversation he/she and the hearer have talked about Mr. Tanaka, he/she uses N_1 *to yuu* N_2, *Tanaka-san to yuu hito* 'a person called Mr. Tanaka'. The state of lack of mutual recognition includes the case where the speaker is not sure whether or not he/she shared the experience of process of recognition in the past. Also, even when the speaker knows that he/she and the hearer can recognize the entity, the speaker uses N_1 *to yuu* N_2 if the speaker does not believe that the speaker and the hearer shared the experience of recognizing the entity prior to the conversation. Thus, the speaker uses N_1 *to yuu* N_2 if the speaker does not believe that the hearer knows that the speaker can recognize the entity. The state of mutual recognition is one of the conditions that determine whether the state of mutual knowledge is symmetrical or not. In that sense, the notion of mutual recognition is parallel to Takubo's theory of discourse management.

3.1. Introduction of a proper noun and the speaker's judgment

When the speaker believes that he/she and the hearer have not shared the process of recognition of the entity, the speaker uses N_1 *to yuu* N_2 to introduce a new proper noun as in (8). This is a case of lack of mutual recognition.

(8) *Watasi wa kinoo **Tanaka-san to/tte yuu hito** ni atta.*
I TOP yesterday Tanaka-Mr. QP say person DAT meet-PERFECT
'I met **a person called Mr. Tanaka** yesterday.'

If the speaker's judgment is correct, that is, if he/she and the hearer actually have not shared the experience of recognizing Mr. Tanaka, the hearer accepts the use of N_1 *to yuu* N_2.

In contrast, (9) and (10) the cases where the speaker misjudged the state of the mutual recognition. In (9), the speaker uses *Tanaka-san to yuu hito* to refer to a person even though he/she and the hearer already shared the process of recognizing the person prior to the conversation. This might cause the hearer to respond in such a manner that he/she questions the speaker's judgment as in B in (9).

(9) A: *Kinoo **Tanaka-san to/tte yuu hito** ni atta.*
yesterday Tanaka-Mr. QP say person DAT meet-PERFECT
'[I] met **a person called Mr. Tanaka** yesterday.'

B: *Tanaka-san, sitte ru yo.*
Tanaka-Mr. know-PROGRESS FP
'[I] know Mr. Tanaka, I tell you.'

In (10), the speaker uses a bare proper noun, Tanaka-san, to refer to a person even though he/she and the hearer have not shared the process of recognition. This might cause the hearer to ask for more explanation about Mr. Tanaka as in B in (10).

(10) A: *Kinoo* ***Tanaka-san*** *ni atta.*
yesterday Tanaka-Mr. DAT meet-PERFECT
'[I] met **Mr. Tanaka** yesterday.'
B: *Tanaka-san. Dare sore?*
Tanaka-Mr. who that
'Mr. Tanaka. Who is that?'

3.2. Lack of mutual recognition and asymmetry in the state of mutual knowledge

In the following example, Takubo's theory of discourse management describes the situation better than Kamio's theory of territory of information. In (11), the speaker is going to talk about Mr. Tanaka whom the speaker met at a party where he/she found out that Mr. Tanaka is an acquaintance of the hearer. Although the speaker knows that the hearer can recognize Mr. Tanaka, the speaker uses N_1 *to yuu* N_2 because the speaker believes that the hearer does not know that the speaker was acquainted with Mr. Tanaka, which is a case of lack of mutual recognition.

(11) A: *Kinoo* ***Tanaka-san to/tte yuu hito*** *ni atta.*
yesterday Tanaka-Mr. QP say person DAT meet-PERFECT
Kimi no siriaina n da tte?
you GEN acquaintance EXT QP

'[I] met **a person called Mr. Tanaka** yesterday. [I] heard it's that [he] is your acquaintance.'

This situation can be described as asymmetry in the state of mutual knowledge by Takubo (1989), and Takubo and Kinsui (1997). They argue that the speaker uses N_1 *to yuu* N_2 in the situation where the speaker's knowledge does not agree with what the hearer thinks about the speaker's knowledge. In other words, the speaker and the hearer have not shared the process of recognizing Mr. Tanaka prior to the conversation.

3.3. Lack of mutual recognition and Constraint on Information Transfer

The theory of discourse management also explains the following example well. Takubo and Kinsui's (1997:749) Constraint on Information Transfer hinders

B in the following dialogue (12), which is same as (6), from using a bare proper noun. In this dialogue, speaker B, who heard about Mr. Tanaka for the first time from speaker A, is processing the information in I-domain. During this processing period, the information cannot be transferred to D-domain; therefore, B continues referring to Mr. Tanaka as *Tanaka to yuu hito* 'a person called Mr. Tanaka'. I describe this period as a period of sharing the process of recognizing the entity. After this period, the state of mutual recognition is established.

(12) A: *Boku no yuuzin ni* ***Tanaka to yuu yatu*** *ga imasu.*-----
I GEN friends among Tanaka QP say person SUB exist
'Among my friends, there is **a person called Tanaka.**'
B: ----*Zya, sono* ***Tanaka to yuu hito*** *ni tanonde kudasai.*
then the Tanaka QP say person DAT ask-REQUEST
'Then please ask **the person called Tanaka.**'
A: *Zya,* ***Tanaka-kun*** *ni tanomimasu.*
Then Tanaka-Mr. DAT ask
'Then [I] will ask **Mr. Tanaka.**'
(Takubo 1989:219, modified)

4. Preemptive usage: use of N_1 *to yuu* N_2 as a strategy

In this section, I examine the case where the speaker uses N_1 *to yuu* N_2 to refer to an entity that the speaker and the hearer can both recognize. I demonstrate that when the speaker uses N_1 *to yuu* N_2 knowing that both the speaker and the hearer can recognize the entity, the speaker uses N_1 *to yuu* N_2 as a strategy to preempt conflict. The speaker uses this strategy when he/she expresses his/her subjective opinion to make the opinion sound objective or when he/she describes his/her action to downgrade the action. By using N_1 *to yuu* N_2, the speaker creates a sense of lack of mutual recognition the conventional usage implicates, which results in increasing the distance between the hearer and the entity. I demonstrate that the sense of lack of mutual recognition N_1 *to yuu* N_2 conventionally implicates plays a significant role in explaining the mechanism of preemptive usage.

4.1. Preemptive usage: use of N_1 *to yuu* N_2 to make the speaker's subjective opinion sound objective

The following are examples where the speaker uses N_1 *to yuu* N_2 to increase the distance between the entity and the hearer to make her opinion sound objective. By doing so, the speaker preempts the hearer's disagreement.

Both examples are from my conversation data. The data was videotaped at a workshop for Japanese teachers in Japan. The speaker in (13) is talking about Japanese people and culture, and the speaker in (14) is talking about education. Both speakers are participating in the group discussion where all participants met for the first time at the workshop. For both examples, the speaker uses N_1 *to yuu* N_2 to refer to a recognizable entity.

(13) *Demo zentaitekini ano* (.)***nihon no:*** (.)***hito tte yuu no*** *wa*
but generally uhm Japan GEN people QP say the one TOP
hakkiri iwanai:desu yo ne.[3]
clearly say-NEG FP FP

'But generally, uhm, (.) **the ones called Japanese** (.) **people** don't say (things) clearly, right?'

(14) *N: demo nanka koo:* (1.0) ***kyooiku*** *toka* ***tte yuu no*** *wa* (.)
uh but somehow uhm education or QP say the one TOP
hitotu no zikan de kanketus-areru mono zyanakutte (.)
one GEN time at complete-PASS thing CP-NEG-and
yappari gakusyuusya ni totte mo soko de kanketu-suru
as you expect learner for also there at complete
mono zyanakute ---
thing CP-NEG-and

'Uh, but somehow, uhm, (1.0) **the one called education** or (something) is (.) not the thing completed at one time, and, (.) as you expect, for learners also, it is not the thing that completes there, and ---.'

Since the workshop is for Japanese teachers in Japan, the entity, Japanese people in (13) and education in (14), should be recognizable to the speaker and the hearer. For these common entities, it is assumed that the speaker and the hearers had a chance to recognize them at some point in the past; thus, the state of mutual recognition has been established prior to the conversation. However, the speakers used N_1 *to yuu* N_2, *nihon no hito tte yuu no* 'the ones (people) called Japanese people' and *kyooiku tte yuu no* 'the one (thing) called education'. The speakers could have used a bare noun, *nihon no hito* 'Japanese people' and *kyooiku* 'education' based on the recognizablity of the entity.

Since the speakers had a choice of using a bare noun, the use of N_1 *to yuu* N_2 is their strategy to induce the sense of lack of mutual recognition, which is conventionally implicated by the linguistic form. The speakers used N_1 *to yuu* N_2 to induce the sense of lack of mutual recognition, which N_1 *to yuu* N_2 conventionally implicates, to increase the distance between the hearer and the entity. The sense of lack of mutual recognition conveys the message that the speaker and the hearer have not mutually recognized the entity, which increases the distance between the hearer and the entity. Through the sense of lack of mutual recognition, the speaker aims to avoid making her point too clear. By

[3] In the excerpts from the conversation data, colons indicate prolongation of vowels, periods in parentheses indicate short pauses, and the numbers in parentheses show the approximate number of seconds of pause.

doing so, the speaker makes her opinion sound objective to preempt the hearer's disagreement.

In both examples, there are some indications that the speaker hesitates to make her statement clear cut. Both speakers pause, prolong vowels, and use fillers, such as *ano:* 'uhm', *n:* 'uh', and *ko:* 'uhm'. These behaviors also show that the speaker intends to avoid making a clear statement. To the speaker, the topic is broad and difficult to discuss, and she knows that the other participants, who are also Japanese teachers, might have different opinions, which motivates her to use the strategy.

Although Suzuki claims that the use of *tte* and *tte yuu no wa* reflects the distance between the speaker and the entity, the mechanism is not clear. For the following example (15), which is same as (7), Suzuki claims that *kimi* 'you' is in the hearer's territory, and the speaker uses the metareferential form, *tte*. This does not explain the mechanism of reflecting the distance between the speaker and the entity, or the speaker's goal in the interaction. In contrast, based on the claim that the form conventionally implicates the lack of mutual recognition, I argue that *tte* in (15) induces the sense of lack of mutual recognition, which results in making the statement sound more objective, which is the speaker's strategy to avoid the hearer's disagreement.

(15)	*Fushigina*	*onna*	*da*	*naa,*	***kimi***	***tte***	***onnna***	*wa.*
	Strange	woman	CP	FP	you	QP	woman	TOP

'you are a strange woman, you (lit. **a woman called you**).'
(Suzuki 1998:443, cited from Martin 1975:1017, modified)

4.2. Preemptive usage: use of N_1 *to yuu* N_2 to downgrade the speaker's action

The followings are examples where the speaker uses N_1 *to yuu* N_2 to increase the distance between the hearer and the entity to downgrade her action. By doing so, the speaker preempts her potential embarrassment.

These examples are also from the same conversation data as I discussed in the preceding section. In (16), the speaker is talking about her occupation, and in (17), the speaker is talking about her thesis. For both examples, the speaker uses N_1 *to yuu* N_2 to refer to the entity that is recognizable to the speaker and the hearer.

(16)	*N: tandaisee*	*(.)*	*ni taisite*	***nihongo*** *(.)*	***to yuu***	*(.)*	***mono*** *o*	
	uh junior college students		to	Japanese	QP say		thing ACC	
	osie:ru	*to yuu*	*koto*	*wa*	*site iru*	*n desu*	*ne.*	
	teach	QP say	thing	at least	doing	EXT	FP	

'Uh, it's that [I']m doing at least the thing called teaching **the thing called Japanese**, you know.'

(17) *Siyootiiki toka* (.)*sirabasu no tigau kyookasyo o* (1.0)
using area or syllabus SUB different textbooks ACC
toriagete (2.0) ***hikaku kentoo to yuu koto*** *o yatta*
investigate-and comparative study QP say thing ACC do-PERFET
n desu kedo
EXT but

'It's that I investigated textbooks that different (geographical) areas are using and whose syllabi are different, and did **the thing called comparative study**, but.'

Like the examples in the preceding section, the speakers in (16) and (17) had a choice to use a bare noun, but they chose to use N_1 *to yuu* N_2 as a strategy. Since the participants of the workshop are Japanese teachers who have the experience of conducting research, both entities, *nihongo* 'Japanese language' in (16) and *hikaku kentoo* 'comparative study' in (17), should be recognizable to the speaker and the hearers. However, the speakers chose to use *nihongo to yuu mono* 'the thing called Japanese language' and *hikaku kentoo to yuu koto* 'the thing called comparative study', respectively. By using N_1 *to yuu* N_2, which conventionally implicates the lack of mutual recognition, the speaker aims to increase the psychological distance between the entity and the hearer, which results in making the action sound vague. The speaker downgraded her action, because she knows that other participants are also teaching and researching Japanese. By downgrading her action, the speaker preempted possible embarrassment.

In (16), the speaker pauses, prolongs vowels, and uses a filler. In (17), the speaker pauses. These hesitations, like the preceding examples, demonstrate that the speakers avoid making their statement clear.

4.3. Hearer's cognitive process: preliminary explanation

Although the purpose of this study is describing the use of N_1 *to yuu* N_2 as the speaker's strategy, the notion of mutual recognition plays a significant role on the hearer's side, as well as on the speaker's side. The following is a potential explanation.

From the hearer's view point, the following cognitive process is possible. In the preemptive usage, "what is said" is true, but "what N_1 *to yuu* N_2 implicates about the entity" is not true. For instance, in (16), the information that the speaker is teaching the thing called Japanese language is true, that is, "what is said" is true. However, lacking mutual recognition of Japanese language between the speaker and the hearers is not true, that is, "what N_1 *to yuu* N_2 implicates about the entity" is not true. In the sense that the speaker uses the linguistic form that implicates the wrong state of mutual recognition of the entity, the speaker violates Grice's (1975) Cooperative Principle. Because of this violation, the hearer is urged to detect "what the speaker is aiming to do", which is in this case increasing the psychological distance between the hearer and the entity.

Borrowing Fauconnier's (1985) term, the linguistic form, N_1 *to yuu* N_2, plays the role of a "space builder" to form a mental space where the entity expressed with N_1 *to yuu* N_2 is identified as one that the hearer does not share the experience of recognizing with the speaker. In the preemptive usage of N_1 *to yuu* N_2, the identification of the entity expressed with N_1 *to yuu* N_2 is not true, and the hearer is urged to reset the mental space.

4.4. Other usage of *tte*

In addition to the preemptive usage that I have discussed, there is another usage where the speaker uses *tte* knowing that the hearer can recognize the entity as in (18). Unlike the preemptive usage, where the speaker expresses his/her subjective opinion, in (18) the speaker is providing factual information about Mr. Tanaka. Furthermore, although *tte* is derived from N_1 *to yuu* N_2, *tte* in (18) cannot be replaced with N_1 *to yuu* N_2–TOP as in (18)', while *tte* in the preemptive usage can be replaced with N_1 *to yuu* N_2-TOP as in (19). Also this usage of *tte* seems to often occur in the situation where the speaker provides hear-say information. I assume that a different mechanism is involved in this usage, and I have excluded this usage from the present study.

(18) *Tanaka-san* ***tte*** *amerika de umareta-soo da ne.*
Tanaka-Mr QP America in was born-I heard CP FP
'[I] heard Mr. Tanaka was born in America.'

(18)' ? ***Tanaka-san to yuu hito wa*** *amerika de umareta-soo da ne.*
Tanaka-Mr. QP say person TOP America in was born-I heardCP FP
'[I] heard **a person called Mr. Tanaka** was born in America.'

(19) ***Tanaka-san tte / to yuu hito wa*** *henna hito da yo.*
Tanaka-Mr QP/QP say person TOP strange person CP FP
'**A person called Mr. Tanaka** is a strange person, I tell you.'

5. Conclusion

I have introduced the notion of mutual recognition to describe the conventional usage of N_1 *to yuu* N_2, and demonstrated that N_1 *to yuu* N_2 conventionally implicates lack of mutual recognition. Then I have explained that in the situation where the speaker can use a bare noun, the use of N_1 *to yuu* N_2 is the speaker's strategy to increase psychological distance between the hearer and the entity. The speaker utilizes the sense of lack of mutual recognition, which N_1 *to yuu* N_2 conventionally implicates. In situations where the speaker expresses his/her opinion, or the speaker describes his/her action, the speaker uses N_1 *to yuu* N_2 to preempt the hearer's disagreement or the speaker's embarrassment, thus, I refer to this usage as preemptive usage.

To understand what the speaker aims to do by using N_1 *to yuu* N_2 in the interaction, the notion describing what the linguistic form implicates is needed in addition to the notion describing the location or the amount of information. I have demonstrated that the notion of mutual recognition is useful to describe what the speaker is doing by using N_1 *to yuu* N_2 in the interaction.

References

Fauconnier, Gilles. 1985. *Mental spaces: Aspects of meaning construction in natural language*. Cambridge: MIT press.

Grice, H. P. 1975. Logic and conversation. In P. Cole & J. L. Morgan (eds) *Syntax and semantics, Vol. 3: Speech acts*. New York: Academic Press.

Jorden, Eleanor Harz with Mari Noda. 1987. *Japanese: The spoken language, Part 1*. New Haven: Yale University Press.

Kamio, Akio. 1990. *Zyoohoo no nawabari riron: Gengo no kinooteki bunseki (Theory of territory of information: Functional analysis of language)*. Tokyo: Taishuukan.

Kamio, Akio. 1994. The theory of territory of information: The case of Japanese. *Journal of Pragmatics* 21: 67-100.

Martin, Samuel E., 1975. A reference grammar of Japanese. New Haven, CT: Yale University Press.

Miura, Akira. 1974. '*To*' to '*tte*' ('*To*' and '*tte*'). *Nihongo Kyooiku* 24: 23-28.

Suzuki, Satoko. 1996. An analysis of *tte*- and *to yuu no wa*-marked phrases in Japanese discourse. *Word* 47: 321-337.

Suzuki, Satoko. 1998. *Tte* and *nante*: Markers of psychological distance in Japanese conversation. *Journal of Pragmatics* 29: 429-462.

Takubo, Yukinori. 1989. *Meesiku no modalitii* (Noun phrase modality). In Y. Nitta & T. Masuoka (eds) *Nihongo no modalitii (Japanese modality)*. Tokyo: Kurosio Publishers.

Takubo, Yukinori. 1992. *Danwa kanri no riron: Taiwa ni okeru kikite no tisiki ryooiki no yakuwari* (Discourse management: Role of the hearer's territory of information in discourse). *Gengo* 21: 52-58.

Takubo, Yukinori and Satoshi Kinsui. 1997. Discourse management in terms of mental spaces. *Journal of Pragmatics* 28: 741-758.

Michiko Todokoro Buchanan
Department of Asian Languages and Literatures
University of Minnesota
453 Folwell Hall
9 Pleasant St. SE
Minneapolis, MN 55455

buch0119@umn.edu

WANT TO/WANNA : Verbal Polysemy versus Constructional Compositionality

GUILLAUME DESAGULIER
Université Michel de Montaigne-Bordeaux 3

0. Introduction

This paper is based on the assumption that *want to/wanna* has grammaticalized from non-modal semantics and non-auxiliary form in Middle English to an emerging member of present-day modal markers (Krug 2000). There is indeed evidence that *want to* has extended beyond the mere expression of lack and volition to the extent that it is more and more used in advice-giving contexts, which tends to show that modalization is at work :

(1) *You want to be careful.*
(2) *You do not want to appear brash or pushy.* (*The Times*)
(3) *It is 10.30 am, so we'll want to go easy.* (*San Francisco Chronicle*)
(4) *You might not want to take those drugs if you're driving.*

Yet this fact does not obliterate the more central meanings of *want (to)* :

(5) *I want an apple.* (volition)
(6) *I want to eat an apple.* (volition + intention)
(7) *Your shoes want polishing.*[1] (lack, need)

At the same time, *want to* displays signs of primary auxiliarization when used deontically and there is every reason to believe that *wanna* is the phono-graphematic consequence of the phonological reduction, morphological contraction and syntactic univerbation of *want* and *to*. (8) below is a good example of a semantic extension with loss of primary meaning and progressive decategorialization :

(8) *(...)then you want to/wanna take the left lane.*

---

[1] A form restricted to British English almost exclusively. In American English *need* will be used instead of *want*.

It is hard to say that the speaker appeals to the hearer's volition or desire in the driving context. In this situation there may be another path that leads to the final destination and *you need to take the left lane* would be more constraining on the part of the speaker. What motivates the advice-giving use and interpretation of *want to* or *wanna* here is the acknowledgment on the part of the speaker that the hearer (*you*) is his/her own authority. In its modal extensions therefore *want to* undergoes a weakening of its original semantic content in the form of 'semantic bleaching' : the primary, more literal meaning is diluted in the extended set of contexts in which it is used. Moreover, in spoken discourse, owing to the advice-giving function of the whole construction, *want to* is very likely to be pronounced [ˈwɒnə], as a single, non-separable unit behaving like an auxiliary governing a bare infinitive. The infinitive marker can even be said to have become an affix or rather a clitic. The same phenomenon can be observed with *gonna, hafta* and *gotta*. Krug (2000) provides an extensive description of the grammaticalization of what he calls 'emerging modals'. Yet previous research on *want to/wanna* has failed to specify what predisposed it to develop from non-modal semantics in Middle English to an emerging member of the modal markers. This is what I set out to do.

In this paper I will bring together linguistic aspects of the motivations for language change that are generally kept apart. To this aim I will argue that the synchronic variation of *want (to)* reflects its historical developments and that it is possible to bring together the binary understanding of language variation (synchronic and diachronic) within a unified framework that reconciles form and meaning. I will show that the combination of such theoretical frameworks as mental space blending and Construction Grammar allows us to unify a broad range of complex cognitive mechanisms at work in language change and the resulting present-day polysemy.

1. The semantic compositionality of *want (to)/wanna*

Sweetser (1999) resorts to Fauconnier and Turner's 'mental space blending' (Fauconnier 1997, Fauconnier & Turner 1996, 2001) to reassert the centrality of semantic compositionality in language. She establishes a general compositional formula using the compositional semantics of the Adj.-N modification construction as a case study. We are about to see in what way the coexistence of the original uses of *want to* –lack, desire, volition– with more recent deontic extensions can be considered as an example of semantic compositionality. The following paragraphs will show to what extent primary scenes interact with frames, active zones and mental spaces when it comes to construing the semantic complexities generated by the intermediary status of *want to* on the VP↔modal auxiliary gradient.

1.1. Conventional metaphoric mappings

Sweetser (1990) contends that semantic change in general is from content to mental and to 'speech act' meanings, not vice-versa. The same kind of unidirectional mappings have shaped the extensions of *want (to)* over time :

I- Sociophysical world ('content') :

(9) *for **wante** of mylke*[2] *(Helsinki)* → lack

II- Mental world ('reasoning') :

(10) *He **wants** some milk / to drink some milk* → desire, volition

III- Discourse interaction ('speech acting') :

(11) *You don't **want to** drink that milk* [it is sour] → advice-giving

The motivation for those successive mappings is rather easy to delineate in semantic terms. It seems natural for anyone to desire what one lacks or needs and to express that desire in an agentive manner by means of volition. Of course, lack does not always entail desire – one may lack and need a flu vaccine without understandably relishing at the prospect of the vaccination itself. Still, one may reluctantly go to the doctor's and ask for the vaccination, prompted by a future benefit – e.g. being in good health – in which case volition does not have to presuppose desire. Conversely, desire does not necessarily induce volition, for we might well find an object, a person or a situation desirable – a villa on the Riviera, an expensive car, etc. – without actually wanting them. Mappings between a state of mind (desire) and an action or speech act (volition, advice giving) are not automatic even if in any case, *want (to)* constructions involve the prospect of some future benefit. The transition from volition to advice-giving and further to polite command is a matter of pragmatic inference. To be understood properly, (8) above requires pragmatic deduction in the appropriate discourse situation. Here, the speaker's strategy is best explained in terms of the following deduction : "it is in the interest of *you* to take the left lane so *you* should[3] do it". The speaker builds a predication that is presented to the hearer and biases the validation of the predication. Such an occurrence can be validated or not but in a *want to/wanna* construction, the hearer is the only one to be able to carry out the validation eventually.

[2] Examples of the kind for *want of + NP or V-ing* are not limited to archaic examples, cf. *for want of anything else*.

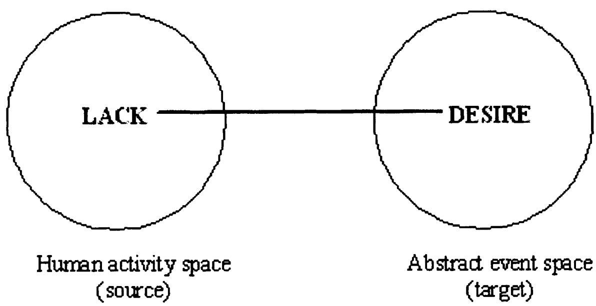
[3] The paraphrase in *should* indicates that we are in a deontic context, more specifically an advice-giving one.

1.2. From lack to desire : primary source and target concepts.

Drawing on Grady's extensive study on primary metaphors and primary scenes (1997) it appears that the transition from lack to desire involves a mapping between a source concept and a target concept. Primary source concepts are "a class of universal, experientially basic properties, relations, actions and processes, which have particular significance with our interactions with the world." (1997:151) whereas target concepts "are more subjective than source concepts, in the sense that they are more strongly tied to internal states." (1997:164)

Lack is a source concept in that it correlates with sensory input from the physical world, while desire is a good instance of a target concept by relating to responses to this input. We can therefore assume that the metaphoric mapping between lack and desire is as shown in Figure 1 below :

Figure 1



It is to be noted that lack is a source concept which also figures in another primary mapping that is often associated with *want to* constructions, except that this time *need* is the target concept. The motivations for these mappings relate to our own experience of lack and to our response to it rather than to its objective manifestations. In other words it is a dynamic primary scene that is involved here.

1.3. Semantic frames

On the other hand, it appears that the concepts of volition and advice-giving encode too much information to be accounted for merely in terms of primary scenes. Accordingly frames seem to be far more promising tools to account for the more complex meaning extensions of *want to*, for "[f]rames are typically represented as collections of related knowledge and experiences, whereas primary scenes are on a much more local scale, and may crosscut many such collections." (Grady 1997:197). As such, "frame structure is a powerful cognitive mechanism which underlies many instances of meaning extension (...)" (Sweetser 1999:134).

We can hypothesize that each subsequent stage in the evolution of *want to* is associated with a frame, such as the Volition Frame and the Advice-Giving

Frame. The former includes a volitional agent (generally in subject position), a desired entity (usually the object) and a hiatus or tension between the two ; the latter consists of some authority reference (rules, laws, etc.) an advice-giver and an advice-receiver.

It is reasonable to think that the primary mappings depicted above are more basic than the Volition Frame, which is in turn more basic than the Advice-Giving Frame.

1.4. Overlap, profiling and active zones.

Even if the conceptualization in terms of a succession of mappings between concepts that are distinct but related gives us a clear idea of the semantic evolution paths of *want to*, one must admit that semantic overlap is the rule once the form appears in discourse.

(12) *American authorities want to take blood samples from thousands of Taliban and al Quaeda prisoners (...)* (*The Times*).

In (12) the volitional property traditionally attached to the verbal notion of *want* is activated. At the same time, identifying the terrorists is presented as an emergency, something that needs to be done, and necessity is a property that should not be ruled out, along with lack – American authorities do not have blood samples – and desire – obtaining blood samples is presented as a future benefit. It is reasonable to think that the different interpretations associated with *want to* (lack, necessity, desire, volition, advice) coexist synchronically in its core meaning. Depending on the context, one or several specific senses will be activated. In (13) below, desire is the predominating activated property :

(13) *I was wanting to leave.*

Contextual interpretation and pragmatic inference will not only profile a given instance of the *want to* constructions as a member of the appropriate frame but also elaborate some active zone[4] of the frame profiled.

More importantly, profiling the Advice-Giving Frame in the case of *want to* is often a matter of argument selection. Thus, an advice-giving usage in discourse very often relies first and foremost on the choice of *you* as the first argument subject :

(14) *If you're looking for a very good collection of tunes to play on your low D whistle, you may want to take a look at this collection of 110 slow airs.* (*www.chiffandfipple.com*)

[4] Following Sweetser (1999), who gives the concept an expanded definition based on Langacker (1987, 1991), 'active zone' subsumes "not only parts or aspects of the entity itself but parts or aspects of the frames associated with it in the complex context of the particular utterance (...)" (1999:147).

When the arguments are co-referential (ie. the advice-giver = the advice-receiver), *want to* is more volitional :

(14') *(...) I want* [me] *to take a look at this collection (...)*

1.5 Mental spaces

Active zone phenomena can also concern "parts or aspects of the entity in another mental space" (Sweetser, 1999:147). Mental spaces as defined by Gilles Fauconnier and Mark Turner as an abstract mechanism meant at describing the interconnections between separate domains of complex referential structures through connectors within and across spaces.

In light of Fauconnier and Turner's theory of Mental Spaces, I claim that it is possible to account for the link between the synchronic polysemy of *want* (NP/*to*) and its diachronic evolution in terms of mental space blending. I contend that each stage in the semantic development of *want* evokes a space structured by some appropriate frame, and that the synchronic meaning of the whole is a successful blending of the two spaces involved. The nature of the blend will depend on the frame profiled. Thus when the Volition Frame is selected, a use like *She wants to eat an apple* might prompt us to associate the desire for an apple – because *she* does not have an apple to eat – with the agentive realization of that desire – on the ground that eating an apple is perceived by *she* as something good, hence something to benefit from (Figure 2). Of course, the number of zones to be activated from inputs 1 and 2 will depend on the context and the level of information to be conveyed. As shown above, (12) invokes more active zones than (13).

On the other hand, when the Advice-Giving Frame is profiled (Figure 3), a warning such as *you want to/wanna be careful* indicates that the speaker (S) identifies the blend as described in Figure 2 (i.e. input 1 in Fig. 3) with his or her own view on the situation (input 2, Fig. 3). In other words, S perceives that being careful is good for *you*, therefore *you* should be careful.

Input 1 and the blend in Fig. 2 together with the blend in Fig. 3 represent the successive stages in the semantic development of *want to*. Furthermore, blending enables us to represent the synchronic coexistence of those different stages in discourse. Thus the study of *want to* reveals that mental space blending can be a good way of conceptualizing synchrony and diachrony. However the problem with the representation of mental spaces is that they fail to represent morphosyntactic blends[5] and semantic blends within the same structure. This is where Adele Goldberg's Construction Grammar comes into play.

[5] And contracted realizations of WANT TO can be considered as blends at the morphosyntactic level.

Figure 2

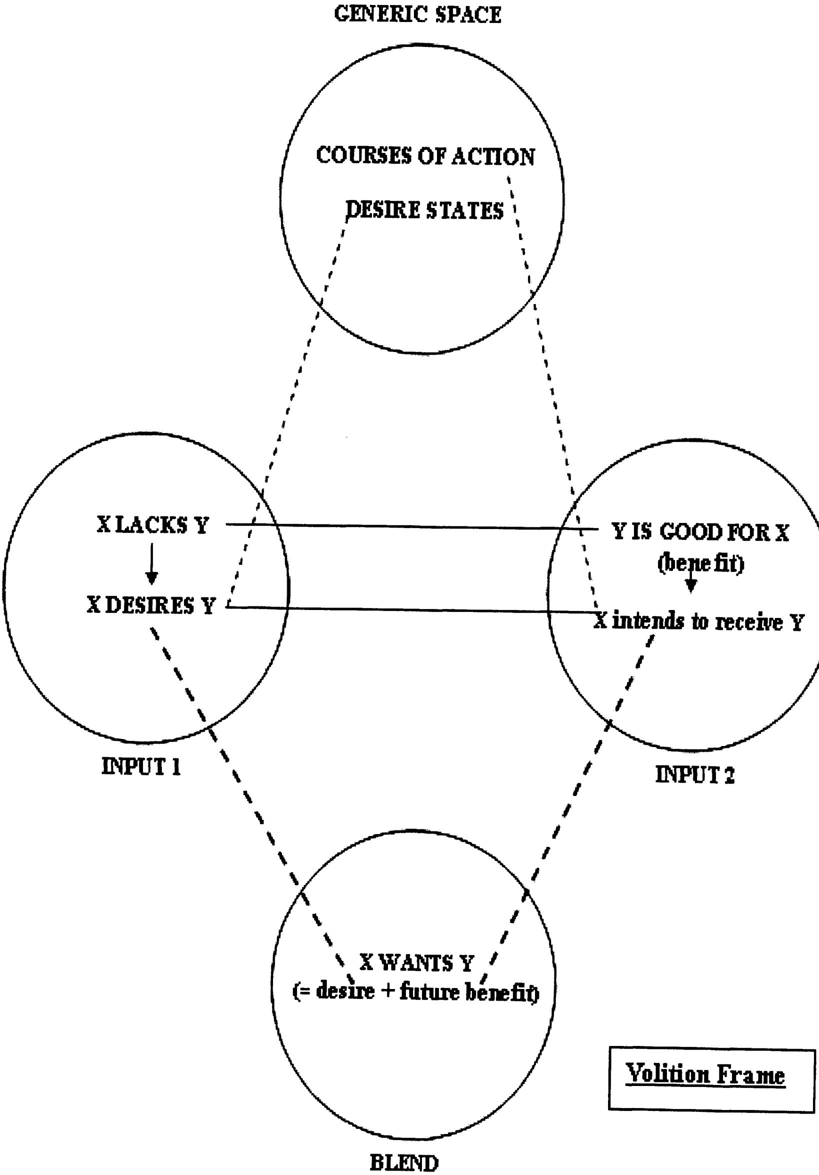
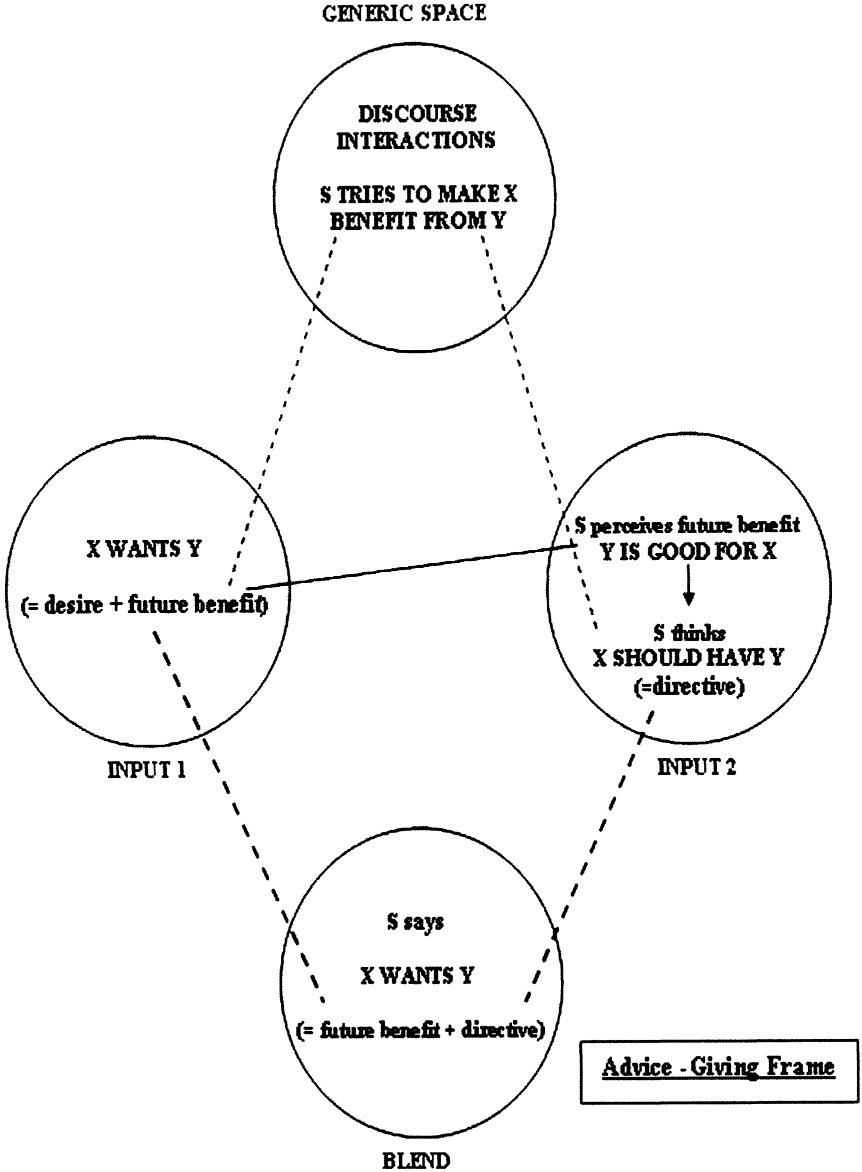

GENERIC SPACE
COURSES OF ACTION
DESIRE STATES
X LACKS Y
X DESIRES Y
INPUT 1
Y IS GOOD FOR X
(benefit)
X intends to receive Y
INPUT 2
X WANTS Y
(= desire + future benefit)
BLEND
Volition Frame

Figure 3


GENERIC SPACE
DISCOURSE INTERACTIONS
S TRIES TO MAKE X BENEFIT FROM Y
X WANTS Y
(= desire + future benefit)
INPUT 1
S perceives future benefit
Y IS GOOD FOR X
S thinks
X SHOULD HAVE Y
(=directive)
INPUT 2
S says
X WANTS Y
(= future benefit + directive)
BLEND
Advice - Giving Frame

2. The constructional compositionality of *want (to)/wanna.*

So far Construction Grammar has proved the best theoretical frame to integrate form and meaning in argument structures. 'Form' in constructions has to do with any combination of syntactic, morphological or prosodic pattern whereas 'meaning' refers to lexical semantics, pragmatics and discourse structure. Drawing on Frame Semantics (Fillmore 1982), Goldberg defines constructions as follows :

> Constructions : pairings of syntax and semantics that can impose particular interpretations on expressions containing verbs which do not themselves lexically entail the given interpretations. (1995:220)

We can posit four constructions which exist independently of the specificity of the verb *want* but which account for its uses :

- WANT (NP/TO) LACK : Potential Recipient lacks Patient → Subj V Obj
- WANT (NP/TO) DESIRE : Potential Recipient desires Patient → Subj V Obj
- WANT (NP/TO) VOLITION : Recipient intends to receive Patient at some future point in time → Subj V Obj
- WANT (NP/TO) ADVICE-GIVING : Agent intends to cause Recipient to receive Patient at some future point in time → [ ] Ag Subj Rec V Obj Pa (or Subj Ag / Rec V Obj Pa if Agent = Recipient)

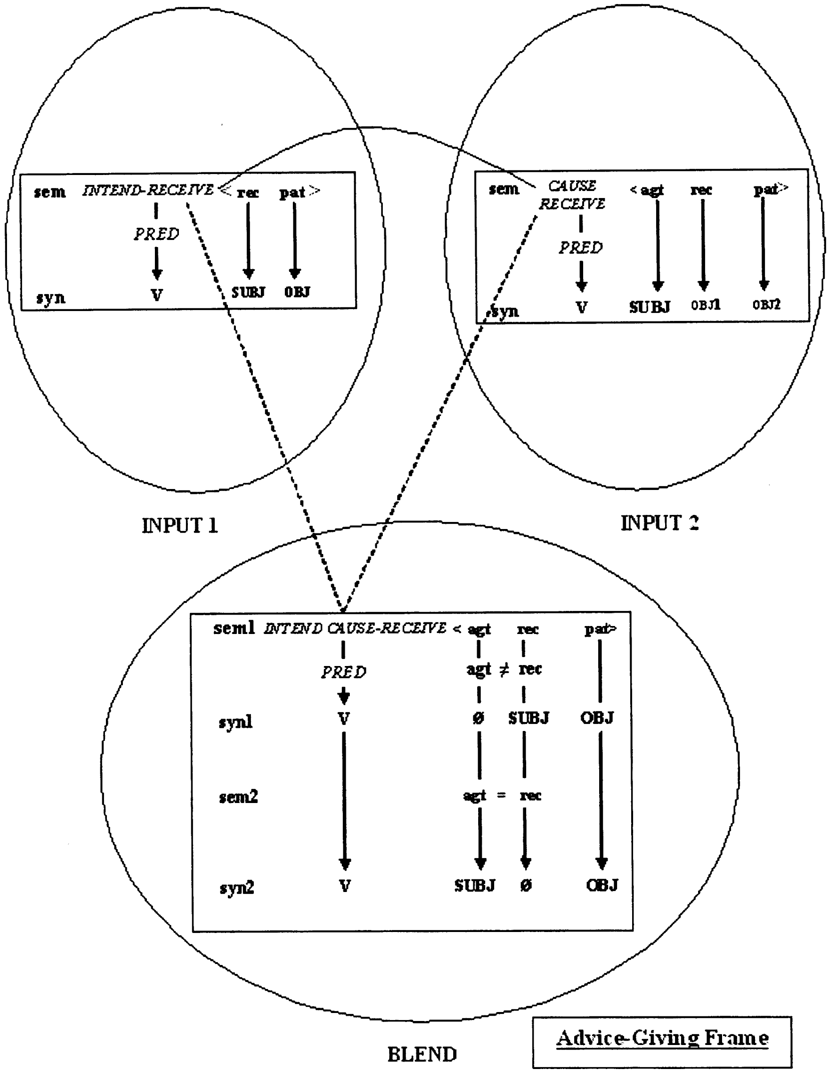
Those constructions are independent from the verbs they account for. For instance, the Volitional Construction also accounts for *want, wish, desire*, etc.

To capture relations among constructions, Goldberg posits asymmetric inheritance links between those which are related both semantically and syntactically : "That is, construction A motivates construction B iff A inherits from B" (1995:72). In other words, for C_2 to inherit from C_1 implies that C_1 dominates C_2 and that C_1 motivates C_2. Those inheritance links apply rather well to the basic constructions associated with *want to* as the Lack Construction dominates and motivates the Desire Construction, which in turn dominates and motivates the Volitional Construction. But the transition from the Volitional Construction to the Advice-Giving Construction is more difficult to handle in terms of inheritance links. For it seems that another construction has to be taken into account, namely the Cause-Receive Construction of which (15) is an illustration.

(15) *I want you to be careful.*

I contend that the Cause-Receive Construction interacts with the Volitional Construction in a compositional way to yield the Advice-Giving Construction (Figure 4).

Figure 4 (the generic space is not analyzed here)



In this respect, *you want to be careful* is interpreted as the successful blending at the constructional level of *you want to be careful* (= intend receive, input 1) and *I want you to be careful* (= intend cause receive, input 2). When the Agent is not the Recipient, the advice-giving use of *want to* is a very convenient and polite way of getting someone to do something without sounding too obtrusive since the Agent does not appear at the surface level, which has the effect of leaving the hearer the choice not to do it, at least at face value. In (3) above, the Agent is also the Recipient (cf. sem2/syn2 in the blend Fig. 4).

This is in keeping with the fact that, owing to its intermediary status, *want to* is increasingly computed as a single construction best analyzed as a "splicing together of a construction associated with its old behaviour and one associated with its new behaviour" (Tabor 1994:170), i.e. a 'hybrid structure', or what I choose to call a 'morpho-syntactic blend'.

Thus advice-giving uses of *want to* prove to be instances of constructional compositionality rather than simple cases of polysemy. The combination of mental space blending and Construction Grammar paves the way for a unified representation of relatively complex cognitive mechanisms by bringing together not only synchronic and diachronic levels of analysis but also form and meaning.

3. Conclusion

The treatment of *want (to)* in the theoretical frame of grammaticalization is the first step towards a unified conceptualization of form and meaning as it shows in what way the structural marks of auxiliarization of *want to/wanna* at the phonological and morphosyntactic levels interact with pragmatic and semantic evidence that the form has been acquiring modal extensions. I have then intended to reassert the centrality of composition in language diversity (synchronically and diachronically) by showing that mental space blending is involved in the auxiliarization and modalization of *want to*. But one major problem with that theoretical construct is that it fails to integrate form and meaning. On the other hand one of the crucial assets of a constructional treatment of *want to* is to integrate form and meaning within the same means of representation. This possibility for *want to* to be used in a wide variety of constructions is precisely what must have predisposed the sequence to be grammaticalized as an emerging modal marker instead of *wish*, *intend* or *desire*. However, this is still a working hypothesis that requires confirmation, notably by analyzing in comparison the constructions associated with *have (got) to/gotta/hafta* and *be going to/gonna*.

Interestingly, we can easily combine mental space blending and Construction Grammar when it comes to showing that the deontic uses of *want to* are best explained in terms of constructional compositionality rather than in terms of polysemy alone. In this respect there is hope for a unified, integrative cognitive approach.

References

Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. Cambridge : Cambridge University Press.

Fauconnier, Gilles and Mark Turner. 1996. Blending as a Central Process of Grammar. In *Conceptual Structure, Discourse, and Language*. Edited by

Adele Goldberg. Stanford : Centre for the Study of Language and Information (CSLI) [distributed by Cambridge University Press]. Pages 113-129.

__________. 2001. Conceptual Integration Networks (Expanded web version, 10 February 2001). http://www.wam.umd.edu/~mturn

Fillmore, Charles J. 1982. Frame Semantics. In : The Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, 111-138. Seoul : Hanshin.

Goldberg, Adele. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago and London : University of Chicago Press.

Grady, Joseph. 1997. Foundations of Meaning : Primary Metaphors and Primary Scenes. PhD dissertation, UC Berkeley.

Krug, Manfred G. 2000. *Emerging English Modals : a corpus-based study of grammaticalization*. Berlin ; New York : Mouton de Gruyter.

Langacker, Ronald. W. 1987. *Foundations of Cognitive Grammar* : Theoretical Prerequisites (Vol. 1). Stanford, Cal. : Stanford University Press.

__________.1991. *Foundations of Cognitive Grammar* : Descriptive Application (Vol. 2). Stanford, Cal. : Stanford University Press.

Sweetser, Eve E. 1990. *From Etymology to Pragmatics*. Cambridge : Cambridge University Press.

__________.1999. Compositionality and blending : semantic composition in a cognitively realistic framework. In Theo Jannsen and Gisela Redeker (eds.), *Cognitive Linguistics : Foundations, Scope and Methodology*. Berlin : Mouton de Gruyter. 129-162.

Tabor, Whitney. 1994. Syntactic innovation : a connectionist model. PhD dissertation, Stanford University.

Guillaume Desagulier
17 rue Camille Desmoulins
75011 Paris, France

gdesagul@etu.u-bordeaux3.fr

Phonological Contrast and Phonetics in Manchu Vowel Systems*

B. ELAN DRESHER and XI ZHANG
University of Toronto

0. Introduction

We argue that the phonological patterning of a phoneme in a grammar is influenced by its contrastive status. A change in this status can lead to a change in its phonological behavior, with little or no outward change in its phonetics. However, the new phonological status of a phoneme may affect its phonetic realizations. Since acquisition of phonological representations is mediated through the phonetics, the proposed analysis posits a mutual interdependence between phonetics and phonology. The synchronic and diachronic patterning of Manchu vowel systems provide evidence in support of this position, as well as for a particular way of understanding contrast.

We will first consider the vowel system of Written Manchu, also called Classical Manchu (Ard 1984, Li 1996) or Literary Manchu (Seong 1989). It is the language of the documents of the Qing (Ching) dynasty (1644-1911) in China. We will then look at diachronic developments that led to the later Manchu dialects, Spoken Manchu (Zhao 1989, Ji et al. 1989) and Xibe (Li and Zhong 1986).

1. The Vowel System of Written Manchu

The vowel system of Written Manchu is shown in (1).

(1) Written Manchu (Zhang 1996)

i		u
		ʊ
— — —	— — —	— — —
	ə	
		ɔ
	a	

Written Manchu has six contrastive vowel phonemes. Given the phonetics of the vowels, we might think that Written Manchu has four or five height classes, but in fact it has only two, as indicated by the horizontal line. There is a set of relatively high vowels above the line and a set of relatively low vowels below the line.

* An earlier version of this paper is Dresher and Zhang 2000. We would like to thank members of the project on Contrast in Phonology at the University of Toronto (http://www.chass.utoronto.ca/~contrast/) for much help over the years. This research was supported in part by grants 410-96-0842 and 410-99-1309 from the Social Sciences and Humanities Research Council of Canada.

Zhang (1996) has argued that the distinction between /u/ ~ /ʊ/ and /ə/ ~ /a/ has to do with the tongue root: the first vowel in each pair is ATR (Advanced Tongue Root), the second is not. ATR vowels tend to be higher than their non-ATR counterparts, which accounts for the difference in height that accompanies the ATR contrast. Li (1996) also proposes that these contrasts are based on the tongue root. He, however, argues that Retracted Tongue Root ([RTR]) is the marked feature that distinguishes the pairs. Zhang (1996) and Zhang and Dresher (2000) provide evidence that [ATR] is marked in Written Manchu.

1.1. ATR Harmony in Written Manchu

All vowels in a word apart from /i/ must agree with respect to ATR. This harmony is most clearly seen in the case of /ə/ and /a/: suffixes with these vowels alternate depending on the ATR value of the stem vowels, as in (2).

(2) ATR harmony in Written Manchu: /ə/ ~ /a/

a.	ATR suffixes			
	həhə	'woman'	həhə-ŋgə	'female'
	susə	'coarse'	susə-də-	'make coarsely'
b.	RTR suffixes			
	aɢa	'rain'	aɢa-ŋɢa	'of rain'
	χʊrχa	'fishing net'	χʊrχa-da-	'catch in a net'

Similarly, /u/ alternates with /ʊ/, as in the suffixes in (3).

(3) ATR harmony in Written Manchu: /u/ ~ /ʊ/

a.	ATR suffixes			
	hərə-	'ladle out'	hərə-ku	'ladle'
	səhəhun	'vertical'	səhə-huri	'towering high'
b.	RTR suffixes			
	baqta-	'contain'	baqta-qʊ	'internal organs'
	laqdaχʊn	'drooping'	laqda-χʊri	'fully drooping'

This latter alternation is apparent only after back consonants, however. The back (dorsal) consonants manifest an allophonic phonetic alternation depending on the [ATR] value of the following vowel: velars [k, g, h] occur before ATR vowels, including [u], and uvulars [q, ɢ, χ] occur before non-ATR vowels, including [ʊ]. In other contexts, /u/ and /ʊ/ merge at the surface into [u], except for a few sporadic examples. This neutralization does not affect the behaviour of /ʊ/ with respect to ATR harmony, as shown in (4). In (4a) [u] derives from /u/, and patterns with ATR vowels; in (4b), [u] derives from /ʊ/, and patterns with non-ATR vowels.

(4) Dual patterning of [u]

a.	ATR harmony with /u/			
	dẓ̜uwə	'two'	dẓ̜uwə-də-	'lean to two sides'
	səktu	'clever'	səktu-kən	'somewhat clever'
b.	RTR vowels with /ʊ/			
	dulba	'careless'	dulba-da-	'act carelessly'
	datṣ̜un	'sharp'	datṣ̜u-qan	'somewhat sharp'

The vowel /i/, though phonetically ATR, is neutral when in position to undergo harmony, as shown in (5).

(5) ATR harmony in Written Manchu: /i/ is neutral

a. With ATR vowels

itṣə	'new'	itṣə-lə	'make new'
sidərə-	'hobble'	sidərə-shun	'hobbled/lame'
sədẓən	'wagon'	sədẓə-tṣi	'wagoner'

b. With RTR vowels

baqtṣin	'opponent'	baqtṣi-la-	'oppose'
bandẓin	'appearance'	bandẓi-sχʊn	'appearance'
tṣaɢan	'books'	tṣaɢa-tṣi	'clerk'

However, when /i/ is in a position to trigger harmony, it occurs only with non-ATR vowels, as in (6).

(6) Stems with only /i/: Suffixes with non-ATR vowels

a. With /a/ suffix

ili-	'stand'	ili-χa	'stood'
itṣi	'direction'	itṣi-ŋɢa	'having direction'

b. With /ʊ/ suffix

idẓi-	'put in order'	idẓi-sχʊn	'obedient'
sifi-	'stick in the hair'	sifi-qʊ	'hairpin'

There thus appears to be only one /i/ phoneme in Written Manchu. Moreover, this phoneme is not specified as being [ATR], unlike /ə/ and /u/. Why not, given that /i/ is phonetically ATR? Intuitively, we might say it is because /ə/ and /u/ have non-ATR 'partners', whereas /i/ does not, as displayed in (7). But a similar explanation will not suffice to account for labial harmony.

(7) Written Manchu: Sets of [ATR] partners

i		u	+
		ʊ	–
ə	+		
a	–	ɔ	

1.2. Labial Harmony in Written Manchu

Written Manchu also has a process of labial harmony (Zhang 1996, Zhang and Dresher 1996, Walker 2001). A suffix vowel /a/ becomes /ɔ/ if preceded by two successive /ɔ/ vowels (8a). Thus, labial harmony is not triggered by a single short or long /ɔ/ (8b), nor by the high round vowels (8c, d).

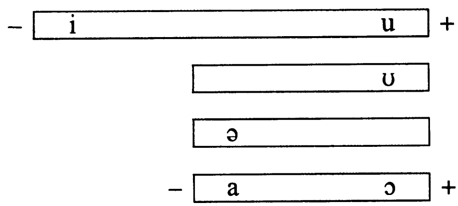
(8) Labial harmony in Written Manchu

a.	bɔtṣɔ	'colour'	bɔtṣɔ-ŋɢɔ	'coloured'
	fɔχɔlɔn	'short'	fɔχɔlɔ-qɔn	'somewhat short'
b.	dɔ-	'alight (birds)'	dɔ-na-	'alight in swarm'
	dɔɔ-	'cross (river)'	dɔɔ-na-	'go to cross'

c. gulu	'plain'	gulu-kən	'somewhat plain'
kumun	'music'	kumu-ŋgə	'noisy'
d. χʊdun	'fast'	χʊdu-qan	'somewhat fast'
dursun	'form'	dursu-ŋɢa	'having form'

Clearly, /ɔ/ and /a/ are 'partners' with respect to the feature [labial]. But why aren't /u/ and /i/? The diagram in (9) appears to parallel (7).

(9) Written Manchu: Sets of [labial] partners



We will argue that it has something to do with /i/ being [coronal]; the evidence is that it causes palatalization of consonants. The specification of /i/ for [coronal] prevents the nonlow vowels from being phonologically specified for [labial], for reasons we will make clear.

1.3. Phonological Feature Specifications of Written Manchu Vowels

Our brief survey of the phonology of Written Manchu vowels has revealed that we require four vowel features to be active in the phonology: a height feature, which we will call [low]; two place features, [coronal] and [labial]; and a tongue root feature, [ATR]. We have also found positive evidence from palatalization that /i/ is specified for [coronal], and from ATR harmony that it must not be specified for [ATR], despite its surface phonetics. In the case of the high back vowels, there is evidence from silence that they are not specified for [labial], again belying their surface phonetics, and in contrast to /ɔ/, which must be specified [labial] in order to trigger labial harmony.

Putting these observations together, we have found positive evidence for the following feature specifications of Written Manchu vowels:

(10) Feature matrix for Written Manchu vowels: Privative features

Phoneme	/i/	/u/	/ʊ/	/ə/	/a/	/ɔ/
[low]				√	√	√
[coronal]	√					
[labial]						√
[ATR]		√		√		

The representations in (10) are underspecified in two different ways. The features in (10) are privative; that is, they have one value (the marked value). Underspecification of this kind is not, strictly speaking, underspecification at all, since there are no available values that could be filled in. Privativity, however, does not account for all the blanks in (10), for some marked specifications are also missing: /i/ is not specified for [ATR], and /u/ and /ʊ/ are not specified for [labial]. The motivation for leaving these values blank has so far been empirical: these representations give the best account of the phonological patterning of the

Written Manchu vowels. What we require now is a theory that gives us precisely these specifications. Our claim is that the absence of marked specifications in (10) is bound up with the notion of contrast, and it is to this topic that we now turn.

2. Contrastive Specification by a Hierarchy of Features

The approach we will adopt has roots in early work (Prague School) on contrastive features. The idea is that to determine contrastiveness of features, it is necessary to determine their relative *scope*, or *ordering*. Thus, Trubetzkoy (1939/1969: 102-103) observes that in the Polabian vowel system, a 'certain hierarchy existed' whereby the back ~ front contrast is higher than the rounded ~ unrounded one, the latter being a subclassification of the front vowels. Trubetzkoy's rationale for this analysis is that palatalization in consonants is neutralized before all front vowels and before 'the maximally open vowel a which stood outside the classes of timbre.' Also, the oppositions between back and front vowels are constant, but those between rounded and unrounded vowels of the same height are neutralizable. The vowel system, according to Trubetzkoy's contrastive distinctions, is given in (11), where the feature [back] has wider scope than does [rounded].

(11) Polabian (Trubetzkoy 1969: 102-3): [back] > [rounded]

Front		Back
(Unrounded)	Rounded	
i	ü	u
ê	ö	o
e		ɑ
a		

In another example Trubetzkoy observes that Greek has a bilabial stop /p/ and labiodental fricatives /f v/, and a postdental stop /t/ and interdental fricatives /θ ð/. Is the primary contrast one of stop vs. fricative or of place? Trubetzkoy appeals to 'parallel' relations between stops and fricatives at different places. In the sibilant and dorsal series (/ts s z/ and /k x ɣ/, respectively), the contrast is unambiguously one of stop versus fricative, since stops and fricatives occur at exactly the same place of articulation. By parallelism, he proposes that the same contrast should apply to the ambiguous cases, which leads to the conclusion that the minor place splits are phonologically irrelevant. The inventory that emerges is given in (12).

(12) Greek: major place, voicing, occlusion > minor place

	Labial		Apical			Dorsal
	bilabial	labiodental	interdental	postdental	sibilant	dorsal
voiceless stops	p			t	ts	k
voiceless fricatives		f	θ		s	x
voiced fricatives		v	ð		z	ɣ

In French, however, Trubetzkoy (1939/1969:126) argues for a split labial series. 'For in the entire French consonant system there is not a single phoneme pair in which the relation spirant : occlusive would occur in its pure form.' Following this analysis to its logical conclusion (n. 93), he disputes that there is an opposition between occlusives and spirants in French because degree of occlusion cannot be regarded independently of position of articulation. The chart in (13) is based on one given by Martinet (1964), who adopts Trubetzkoy's analysis.

(13) French obstruents (based on Martinet 1964:65)

	bilabial	labio-dental	apical	alveolar	pre-palatal	dorso-velar
voiceless	p	f	t	s	š	k
voiced	b	v	d	z	ž	g

Thus, Greek and French require a different ordering of the continuant feature relative to minor place features. Trubetzkoy's discussion implies that place features take scope over occlusion (French) unless an occlusion contrast is needed anyway (parallelism, Greek).

The contrastive hierarchy had its heyday in the 1950s, when it was proposed by Jakobson and Halle in a number of publications, including *Fundamentals of Language* (1956) and Halle's *Sound Pattern of Russian* (1959). Jakobson and Halle (1956) refer to this hierarchy as the 'dichotomous scale', and adduce 'several weighty arguments' in support of this hierarchical approach to feature specification. One argument involves language acquisition. They propose that the order of these contrastive splits is partially fixed, thereby allowing for certain developmental sequences and ruling out others. The sequence in (14), for example, concerns oral resonance (primary and secondary place) features.

(14) Predicted acquisition sequences (Jakobson and Halle 1956: 41)

dental vs. labial consonants

narrow vs. wide vowels

palatal vs. velar narrow vowels

velopalatal vs. labial and dental consonants

pal vs. vel wide Vs

rnd vs. unrnd narrow pal Vs

rnd vs. unrnd wide pal Vs

unrnd vs. rnd velar Vs

pal vs. vel Cs

rnd vs. unrnd or pharyn vs. nonpharyn Cs

pal vs. nonpal Cs

See Rice and Avery (1995) for a more recent application of this idea to the acquisition of representations. On this view, a natural way of determining contrast is by splitting the inventory by means of successive divisions, governed by an ordering of features. An algorithm corresponding to this idea, which we call the Successive Division Algorithm (SDA), is given in (15) (Dresher 1998a, b, 2002, 2003, and Mackenzie and Dresher this volume).

(15) Successive Division Algorithm

a. In the initial state, all sounds are assumed to be variants of a single phoneme.

b. If the set is found to have more than one phoneme, a binary distinction is made on the basis of one of the universal set of distinctive features; this cut divides the inventory into a marked set and an unmarked set. The selected feature is *contrastive* for all members of these sets.

c. Repeat step (b) in each set with the next feature in the hierarchy, dividing each remaining set until all distinctive sounds have been differentiated.

d. If a feature has not been designated as contrastive for a phoneme, then it is *redundant* for that phoneme.

We assume that the ordering of features can vary within limits; to the extent that variation is permitted, learners must fix the order based on language-particular phonological and phonetic evidence (see the papers in Hall 2003 for further recent studies in contrast from this perspective).

3. The Contrastive Hierarchy of the Written Manchu Vowel System

The patterning of the Written Manchu vowel system is consistent with the contrastive hierarchy in (16).

(16) Written Manchu contrastive hierarchy
[low] > [coronal] > [labial] > [ATR]

The fact that labial harmony is confined to the low vowels suggests that the height contrast is more fundamental than the labial contrast: labial harmony operates within a domain defined by the feature [low]. This conclusion is supported by the evidence that the high vowels have no contrastive specification for [labial]. In terms of the SDA, this suggests that the feature [low] has wider scope than [labial].

Similarly, we can establish that [coronal] must take precedence over [labial]. For if [labial] > [coronal], then the nonlow back vowels would be designated as [labial], and /i/ would not require a specification for [coronal]. The ordering [coronal] > [labial] gives us the desired specifications.

Evidently, the feature [coronal] is not contrastive among the [low] vowels because none of these vowels meet the requirements for coronality. Therefore, [coronal] fails to make any contrasts in the [low] region, and so the way is open for [labial], the next feature in the hierarchy, to be assigned to /ɔ/.

We have observed that /i/ has no phonological [ATR] feature. Since /i/ is phonetically [ATR], its lack of such a specification in the phonology must be due, on the approach taken here, to its contrastive status. We have also seen that /i/ is specified for [coronal]. Considering the relative scopes of these features, it must be the case that [coronal] takes scope over [ATR]. For then, /i/, which is the only [coronal] vowel, would already be distinguished from all other vowels, and so would not be eligible to receive any further specifications. On this ordering, the feature [ATR] is needed in the nonlow vowels only to distinguish /u/ from /ʊ/.

The sets of [ATR] partners in (7) suggest that [labial] > [ATR]. For then [ATR] is relevant only to /ə/ and /a/ among the [low] vowels, and not to /ɔ/.

Finally, because Written Manchu has only one potentially [coronal] vowel, /i/, we obtain the same results with either [low] or [coronal] taking precedence.

Zhang has observed that a two-height system is very stable across all the Manchu-Tungus languages surveyed in Zhang 1996, suggesting that the division into two height classes is a basic property of these vowel systems. This would argue in favour of ordering [low] highest. Thus, we arrive at the Written Manchu contrastive hierarchy shown in (16).

4. The Evolution of the Manchu Vowel System

Subsequent changes in the Manchu vowel system support our analysis of Written Manchu and the general approach to contrast it is based on, and demonstrate the interplay between phonological patterning and phonetics. The vowels /ə/ and /u/ will undergo changes in their phonological representations and behaviour without any big phonetic change at first. We will show how their ambiguous status leads learners to make changes in the grammar that later influence their phonetics.

Spoken Manchu (Zhao 1989, Ji et al. 1989) is a later form of Written Manchu. It displays some interesting continuities with the older form of the language, as well as some striking differences that shed further light on the role of contrast in phonology. The vowel system of Spoken Manchu is presented in (17).

(17) Spoken Manchu (based on Zhao 1989, Ji et al. 1989)

i	y	ə	u
ɛ		a	ɔ

In comparing this vowel system with that of Written Manchu in (1), we note a number of differences, which we tabulate in (18). We will argue that change (18a) sets the stage for (18b), and that (18c) follows from (18b).

(18) Differences between Written Manchu and Spoken Manchu vowel systems
   a. Spoken Manchu no longer has a contrast between /u/ and /ʊ/.
   b. In Spoken Manchu /ə/ is a nonlow vowel, in Written Manchu it is a low ATR counterpart to /a/.
   c. Spoken Manchu has added coronal phonemes /y/ and /ɛ/.

We will begin with the first change. We observed that in Written Manchu the contrast between /u/ and /ʊ/ was already neutralized to [u] in most contexts, surviving only after velar/uvular consonants and sporadically in other contexts. It is no surprise, therefore, to see this neutralization continue to completion in Spoken Manchu, resulting in the total merger of /u/ and /ʊ/ into [u] and the complete loss of the /ʊ/ phoneme.

In a contrast-driven approach, the loss of a contrast in one part of the vowel system could have wider effects. In the Written Manchu system, the /u/ ~ /ʊ/ contrast involves the feature [ATR], just like the /ə/ ~ /a/ contrast. The unity of the [ATR] contrast is made more salient by the rule of ATR harmony, which clearly indicates to language learners that the vowels are to be sorted into [ATR] and non-ATR sets. But with the loss of /ʊ/, the position of [ATR] in the system becomes much more tenuous. The vowel /u/ would now join /i/ as a neutral vowel, occurring with both [ATR] and non-ATR vowels.

Now, the only evidence for an [ATR] contrast falls on /ə/ ~ /a/. Many languages have these vowels in their inventories without having a contrastive feature [ATR]. As we observed earlier, the contrast between these vowels could more straightforwardly be attributed to a difference in height. Indeed, the feature [low], which is required independently, can also serve to distinguish /ə/ from /a/.

Therefore, without assuming that the phoneme /ə/ changed phonetically, the loss of /ʊ/ could have indirectly led to a change in the phonological status of /ə/, from [low] to nonlow. This reclassification, in turn, could have influenced the phonetic realizations of /ə/, because in Spoken Manchu it is definitely a nonlow vowel. Zhao (1989) characterizes it as a mid-high back unrounded vowel, with an allophone [ɣ]; according to Ji et al. (1989), [ə] is in free variation with a high back unrounded vowel [ɯ]. It is reasonable to suppose that there is a mutual influence between phonology and phonetics in such cases. The phonetics of a vowel obviously influence its phonological representation; but this influence is not simply one way, and the phonological representation can in turn affect the phonetics, by defining the space within which the vowel can vary.

The change in status of /ə/ in turn has consequences for the specification of /u/. Recall that in Written Manchu we found evidence that the vowel /i/ is phonologically [coronal], but no evidence that the vowels /u/ and /ʊ/ are phonologically [labial], though they are phonetically round. Recall also that this lack of [labial] specification is entirely expected under the theory of contrastive specification we are assuming: because only a single place contrast exists in the nonlow vowels, that contrast can be either [coronal] or [labial], but not both.

The elevation of /ə/ to a nonlow vowel, joining /i/ and /u/, changes the situation. Assuming, as before, that [coronal] takes precedence, /i/ is again specified [coronal], distinguishing it from /ə/ and /u/. But now we must still distinguish the latter two vowels from each other. The most straightforward distinction is again a place distinction, whereby /u/ is specified [labial], as diagrammed in (19).

(19) Spoken Manchu after loss of /ʊ/

[coronal]		[labial]	
i	ə	u	
	a	ɔ	[low]

There is positive evidence that Spoken Manchu /u/ has acquired a [labial] specification. We cannot appeal to labial harmony, because both labial and ATR harmony have been destroyed in Spoken Manchu (Zhang 1996). However, the development of the new phonemes /ɛ/ and /y/ does provide evidence bearing on this question. According to Zhang (1996), Spoken Manchu /ɛ/ often corresponds to Written Manchu /a/ when followed by /i/. It is likely, then, that this phoneme originated from /a/ followed by /i/. Since /i/ has a [coronal] feature and /a/ has a [low] feature, it follows that the addition of the [coronal] feature from /i/ to an /a/ would result in a [low, coronal] vowel, namely /ɛ/. This development could have begun even in Written Manchu, since the features that participate in the process were all in place. Over time, however, as various other changes caused the original environment of the rule to become obscure, the vowel /ɛ/ started appearing in unpredictable contexts and became a new phoneme.

The Spoken Manchu vowel /y/ also developed from a sequence of vowels. As Zhang (1996) shows, Spoken Manchu /y/ corresponds to Written Manchu /i/ followed by /u/, as well as /u/ followed by /i/. Now, /y/ is a front round vowel and thus has the features [coronal] and [labial]. The feature [coronal] is contributed by /i/, parallel to its role in the creation of /ɛ/. But the feature [labial] must come from /u/. In Written Manchu, we have argued that this vowel did not possess a [labial] feature, but that in Spoken Manchu, following the elevation of /ə/ to a nonlow vowel, it does. The creation of /y/ thus provides evidence for the [labial] specification of /u/ in Spoken Manchu.

Like /ɛ/, the new vowel /y/ came to stand in environments where it could not simply be analyzed as deriving from /i/ and /u/, and thus became a separate phoneme which does not depend on receiving a [labial] specification from /u/. However, the development of [y] in the first place provides evidence for a labial feature on /u/.

We have seen, then, that the vowel systems of Written Manchu and Spoken Manchu act as expected given our theory of contrastive specification. Further evidence supporting this approach comes from Xibe (Sibo), another descendent of Written Manchu. The development of the Xibe vowel system is similar to that of Spoken Manchu: the contrast between /u/ and /ʊ/ has been lost along with the feature [ATR], the vowel /ə/ has been reinterpreted as a nonlow vowel, and new phonemes /y/ and /ɛ/ have developed from combinations of other vowels. As in Spoken Manchu, the development of these new phonemes supports the theory that /u/ has acquired a [labial] specification. In addition, a third new vowel, /œ/, has arisen, most likely from earlier /ɔ/ followed by /i/ (Zhang 1996:126).

Unlike Spoken Manchu, Xibe retains a labial harmony rule. According to Li and Zhong (1986), there is an alternation between /ə/ and /u/ in Xibe suffixes (this alternation is not found in Norman 1974). /u/ occurs if the stem-final vowel is round (20a), /ə/ occurs otherwise (20b).

(20) Alternation between /ə/ and /u/ in Xibe suffixes

	Written Manchu	*Xibe*	*Gloss*
a.	bu-hə	bu-xu	'gave'
	bɔdɔ-χɔ	bɔdu-χu	'thought'
	hətu-kən	xətu-kun	'somewhat stocky'
	farχʊ-qan	farχu-qun	'somewhat dark'
	fɔχɔlɔ-qɔn	fœχulu-qun	'somewhat short'
b.	gətə-hə	gətə-xə	'awoke'
	ana-χa	anə-χə	'pushed'
	nətʂi-kən	nətɕi-kən	'somewhat flat'
	amba-qan	am-qən	'somewhat big'

Recall that in Written Manchu labial harmony was restricted to the low vowels, and created an alternation between /a/ and /ɔ/. In Xibe, noninitial vowels tended to be raised – almost always in suffixes, frequently in stem vowels – so an original sequence of the form /a/ - /a/ would become /a/ - /ə/ or /ə/ - /ə/, and a sequence of the form /ɔ/ - /ɔ/ would become /ɔ/ - /u/ or /u/ - /u/. The labial harmony observed in Xibe is not merely a holdover of Written Manchu labial harmony, however, for in Xibe harmony is triggered not only by /u/ derived from older /ɔ/, but also by original /u/. The fact that /u/ triggers and undergoes labial harmony further supports the hypothesis that it has a [labial] specification in Xibe.

5. Conclusion

We have argued that an approach to contrastive specification in terms of the contrastive hierarchy provides an illuminating account of the vowel system of Written Manchu, and of the evolution of the later Manchu languages. In this approach, features are specified following the Successive Division Algorithm and a hierarchy of features. In Manchu, the hierarchy has the feature [low] at the top, followed by [coronal], then [labial], and finally [ATR] at the bottom. The analysis also illustrates the subtle interplay between phonetics and phonological contrast.

References

Ard, Josh 1984. Vowel harmony in Manchu: A critical overview. *Journal of Linguistics* 20: 57-80.

Dresher, B. Elan. 1998a. Child phonology, learnability, and phonological theory. In Tej Bhatia and William C. Ritchie (eds.) *Handbook of Language Acquisition*, 299-346. New York: Academic Press.

Dresher, B. Elan. 1998b. On contrast and redundancy. Paper presented at the annual meeting of the Canadian Linguistic Association, Ottawa. Ms., University of Toronto.

Dresher, B. Elan. 2002. The contrastive hierarchy in phonology. Paper presented at the Second International Conference on Contrast in Phonology, Toronto. In Hall (2003).

Dresher, B. Elan. 2003. Contrast and asymmetries in inventories. In Anna-Maria di Sciullo (ed.) *Asymmetry in Grammar, Volume 2: Morphology, Phonology, Acquisition,* 239-257. Amsterdam: John Benjamins.

Dresher, B. Elan and Xi Zhang. 2000. Contrast In Manchu vowel systems. Paper presented at the 1st International Conference on Manchu-Tungus Studies (ICMTS 1), University of Bonn, August 28 – September 1, 2000.

Hall, Daniel Currie (ed.). 2003. *Toronto Working Papers in Linguistics* 20 (*Special Issue on Contrast in Phonology).* Toronto: Department of Linguistics, University of Toronto.

Halle, Morris. 1959. *The Sound Pattern of Russian: A Linguistic and Acoustical Investigation.* The Hague: Mouton. Second printing, 1971.

Jakobson, Roman and Morris Halle. 1956. *Fundamentals of Language.* The Hague: Mouton.

Ji, Yonghai, Zhizhong Zhao and Liyuan Bai. 1989. *Xiandai Manyu Babaiju* [*Eight Hundred Sentences of the Modern Manchu*]. Beijing: Zhongyang Minzu Xueyuan Chubanshe.

Li, Bing. 1996. *Tungusic Vowel Harmony: Description and Analysis* (HIL dissertations 18). Dordrecht: ICG Printing.

Li, Shulan and Qian Zhong. 1986. *Xiboyu Jianzhi* [*A Brief Introduction to the Xibe Language*]. Beijing: Minzu Chubanshe.

Mackenzie, Sara and B. Elan Dresher. This volume. Contrast and phonological activity in the Nez Perce vowel system.

Martinet, André. 1964. *Elements of General Linguistics.* With a Foreword by L. R. Palmer. Translated by Elisabeth Palmer. Chicago: University of Chicago Press.

Norman, Jerry. 1974. A sketch of Sibe morphology. *Central Asiatic Journal* 18: 159-174.

Rice, Keren and Peter Avery. 1995. Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In John Archibald (ed.) *Phonological Acquisition and Phonological Theory*, 23-42. Hillsdale, NJ: Lawrence Erlbaum.

Seong, Baeg-in. 1989. Vowel length in Manchu. *Proceedings of the XXVIII Permanent International Altaistic Conference.* Wiesbaden: Otto Harrassewitz.

Trubetzkoy, N. S. 1939. *Grundzüge der Phonologie.* Göttingen: Vandenhoek & Ruprecht. Translated by Christiane A. M. Baltaxe (1969) as *Principles of Phonology*. Berkeley: University of California Press.
Walker, Rachel. 2001. Round licensing, harmony, and bisyllabic triggers in Altaic. *Natural Language and Linguistic Theory* 19: 827-878.
Zhang, Xi. 1996. Vowel systems of the Manchu-Tungus languages of China. Doctoral dissertation. University of Toronto.
Zhang, Xi and B. Elan Dresher. 1996. Labial harmony in Written Manchu. *Saksaha: A Review of Manchu Studies* 1: 13-24.
Zhang, Xi and B. Elan Dresher. 2000. Tongue root harmony in Written Manchu. Paper presented at the 1st International Conference on Manchu-Tungus Studies (ICMTS 1), University of Bonn, August 28 – September 1, 2000.
Zhao, Jie. 1989. *Xiandai Manyu Yanjiu* [*Studies on the Modern Manchu Language*]. Beijing: Minzu Chubanshe.

Department of Linguistics
University of Toronto
Toronto, Ontario
Canada M5S 3H1

dresher@chass.utoronto.ca
zhang1@chass.utoronto.ca

I bet you think this paper is about *you*: Participant roles and *you*

RANDALL EGGERT
University of Utah

Linguists generally classify *you* as an indexical, thereby assuming that it indexes the addressee, which must be an integral part of the sentence context (cf. Kaplan 1989, Wettstein 1991).[1] Under Kaplan's (1991) Logic of Demonstratives, just as every context contains an agent, time, and position to which tokens of *I*, *now*, and *here* refer respectively, we can presume that every context contains an addressee to which a token of *you* refers. This is how well-behaved indexicals operate. However, indexicals often misbehave (cf. Clark and Carlson 1991, Eggert 1998, Levinson 2000, McCawley 1984, Nunberg 1993). In this paper, I consider cases in which tokens of *you* misbehave, where the referents are not fully determined by the physical contexts of the utterances, as in (1-4).

(1) Some of you know the news; I'm not talking to you; I'm talking to the rest of you.
(Levinson 2000: 178)

(2) Let this be a warning from someone who learned the hard way. The disabled community is different from other minority groups in one major way: You never know when you will join us.
(Zabarsky, M. Our right, your convenience. *Boston Globe*. 1/17/1996, cited in Kleinedler 2000:225)

(3) Unattended vehicles on the drive-through will be impounded. Return to your car immediately.
(Announcement at Sea-Tac Airport, February, 2000)

(4) Asian? Why God made you… This quarter with Asian American students for Christ 7 p.m. Fridays, Ida Noyes W. Lounge.
(Posting, University of Chicago Campus, 1999)

[1] A preliminary version of this paper was presented as a talk for the Linguistics Colloquium at the University of Utah. I would like to thank my colleagues and students for their support and feedback. Naturally, none of them can be held responsible for any shortcomings in the paper.

Levinson (2000) argues that the narrowing of 'you' in the second and third instances in (1) are examples of generalized implicatures. Kleinedler (2000) expands this analysis arguing that the narrowing in (2) involves a particularized implicature. In this paper, I approach such examples not as a neo-Gricean, but as a neo-Goffmanian. I will argue that the addressee is not an entity, but a participant role. A recipient may inhabit the addressee role, but, as I will show, an inhabitant of the role is not automatically the referent of "you". I will then suggest a revised version of Nunberg's (1993) model for indexicality.

Goffman (1976; 1981) suggested that the traditional notions of speaker and addressee are based on a canonical speech situation. In a canonical speech situation, a speech event has exactly two participants who are close to each other both in time and space. At any given moment, one of the participants is speaking, hence is the speaker, and the other is hearing, hence is the addressee. The two participants take turns so that each has the opportunity to be the speaker, and each has the opportunity to be the addressee. However, much—perhaps most—of the time we use language in a situation that is non-canonical. Therefore, Goffman separated speaker and hearer into different roles (though some are not properly participant roles, e.g., the non-ratified recipients).[2]

(5) Sender :
- Animator: 'the individual who produces the words'
- Author: 'the individual who chooses the words and ideas expressed'
- Principal: 'the individual who is responsible for any beliefs expressed, any positions established, and any commitments created by the words'

Recipient:
- Ratified recipient: 'individual who is officially part of the discourse'
  - Addressed recipient: 'individual to whom an utterance is addressed'
  - Non-addressed recipient
- Non-ratified recipient:
  - Eavesdropper: 'individual who intentionally receives a message they are not ratified to receive'
  - Overhearer or bystander: 'individual who accidentally receives a message they are not ratified to receive'

Levinson (1988) extends this list to 17 different roles, 11 of which are participant roles. However, in separating these different roles, we're left with an awkward question: Why is it that cross-linguistically languages only distinguish between 1st, 2nd, and 3rd persons?

[2] In order to expand the range of speech events to include non-spoken language, I have replaced Goffman's terms *speaker* and *hearer* with *sender* and *recipient*.

> If the traditional roles of Speaker and Addressee are so patently oversimplified, why is it that natural languages develop correspondingly simple pronominal systems, instead of more "accurate" ones in which *I* and *you* would be split into a dozen or so distinct forms for the distinct role relations they may signal?
> (Hanks 1996:165)

Hanks (1996) and Irvine (1996) both argue that we only need to posit two participant roles, as long as we properly use another of Goffman's concepts: embedding (cf. Goffman 1974, 1981). Speech events can be embedded inside of other speech events; previous, future, and hypothetical speech events can be indexed by the current speech event. As Hanks and Irvine demonstrate, a given utterance may be connected to more than one speaker because it can embed multiple speech events. Thus, we only need one role for sender. Of course, an utterance may also be connected to more than one addressee, which can help us to reduce to one Levinson's (1988) four recipient participant roles (interlocutor, indirect target, intermediary, and audience). We needn't concern ourselves with any of the other roles, which are not participant roles.

Now, I return to the question of *you*: Who is 'you'? 'You' is evidently associated with the addressee, just as 'I' is associated with the speaker. Thus the question becomes: Who is the addressee? This question is rarely asked in the literature. Yet it seems crucial to understanding how conversations work. In other words, how do we know when we are 'you'?

Hanks (1993) and Silverstein (1995) suggest that indexicals are simultaneously presuppositional and creative. They presuppose aspects of the discourse context, and they establish aspects of it. In particular, Silverstein (1995: 206) includes second person in the class of relatively creative (or performative) indexes, which "function as the signal for the existence of speech-event features, as in the choice of pronominals, which assign the event roles of speaker, hearer, audience, and referent to certain individuals."

Hanks (1993: 141) describes a situation in which a speaker uses *I* and *you* to describe a fictional, hypothetical conversation. The speaker "postulates a deictic framework identical to the current one in all but the ways that he specifies" and "presupposes the intelligibility of our current shared interactive framework, and uses this as a template from which to generate new ones" (141). He goes on to write:

> The transposability of our current interactive framework into an indefinite number of other hypothetical ones is possible just because speakers have a common sense of the typical ways in which frameworks are the same.
> (Hanks 1993: 142)

This "transposing" of interactive frameworks is essentially what Goffman termed *transformation*. He suggests that transformations are a subcategory of embeddings: "We not only embed utterances, we embed interaction arrangements" (Goffman 1981: 153). A good example is a love song, which is a transformation of a more prototypical framework, such as that of a love letter. In a typical love letter, there is one sender and one recipient, the 'I' and the 'you' respectively. In a typical love song, there is one sender and multiple recipients. Generally, the singer is taken to be the 'I' of the song, but the 'you' may not even count among the actual hearers. We understand the song, in part, because we understand the more common framework of a love letter.

Consider, for example, the lyrics from Carly Simon's song:

(6) You're so vain. I bet you think this song is about you.

Because we know the typical framework for such utterances, and because we know the meaning of *you*, we assume that "you" in the song refers to a particular person. This is the presuppositional aspect of the indexical. But the creative aspect of this use of *you*, is more indeterminate. Who is the actual referent? At the time the song was released, there were four candidates: Mick Jagger (who, incidentally sang back up), Kris Kristofferson, Cat Stevens, and Warren Beatty, all of whom Simon had previously dated (www. carlysimon.com/vain.vain/htm, accessed May 2003). In fact, Beatty called her shortly after to thank her for writing about him.

Beatty's call brings up an important aspect in how people use *you*: both the sender and the recipient actively create its referent. Self-selection is more important in determining the referent of a token of *you* than is commonly assumed. In elementary school, my schoolmates played a playground game. A kid would yell: "Hey, you loser!" Inevitably, somebody—say Randy—would turn around to see if he had been addressed, at which point all of the other kids would point at him, saying: "Randy turned around; he must be a loser." Thus is the logic of the playground: self-selection makes you 'you'.

However, the playground ignores half of the equation. In normal situations, reference is a cooperative act that requires at least two participants (cf. Clark and Wilkes-Gibbs 1990 for a model of reference as a cooperative act). So when Beatty called Simon to thank her for the song, she could have told him, "I'm sorry, but you are not 'you'."

Now let's turn to another transformation, a transformation from personal notes to public notes. With a personal note, a sender typically leaves a message for a specific recipient in a place where he or she will find it and nobody else will. For example, a wife might leave a note for her husband on the refrigerator. With a public note, the sender leaves a message in a public space; the sender does not know who or how many people will receive the message. (7-9) are examples.

(7) I want your stuff
Moving? Sell me the stuff you aren't taking with you!
(Posting. University of Chicago, 7/98)

(8) Moving? Too many things to take with you? I will buy your TV (and VCR? or Either one?).
(Posting. University of Chicago, 5/99.)

(9) We want to live in your Apartment! Responsible students looking to sublet 2 or 3 bedroom apartment this summer (June- September or October). (Posting. University of Chicago, 4/99.)

Such public notes are common. But consider how a reader might respond. When I read the first line of (7), in my head I responded: "No you don't." After all, my stuff was not worth wanting. But my response, I realized, was a joke, analogous to a person responding to (10) with a simple affirmative.

(10) Do you know what time it is?

My internal response was intentionally uncooperative. I knew quite well that, despite the fact that I was moving, I was not the referent of "you". But imagine I had been so misguided to think that anybody would want my stuff; had I called the sender of the message, they would have undoubtedly told me that they did not want my stuff. In other words, I was not 'you'.

In such public notes there are linguistic and extra-linguistic clues for who could qualify as a candidate for the referent of "you". In (8), for example, the one-word question in the first line tacitly limits the candidates to those who are moving. In (9), candidates are not only limited to those who have an apartment they wish to sublet, but to those who have desirable apartments. The latter aspect is extralinguistic, but is clearly part of the schema we assume when one person rents an apartment to another (i.e., the renter must desire to live in the apartment).

(11-13) further illustrate this last point.

(12) Get paid to see your friends graduate! ... If you are interested, please call...
(Posting. University of Chicago, 12/98.)

(13) 3 out of 4 med students who took a commercial MCAT prep course took Kaplan. Shouldn't you?
(Posting. University of Chicago, 5/99.)

(14) Don't let Organic Chemistry keep you from getting into med school.
(Posting. University of Chicago, 5/99.)

In (11), there are a number of criteria to qualify as a referent of "you": 'you' must have friends who are graduating from the U of C; 'you' must not be graduating from the U of C; 'you' must not have a ticket to attend graduation; 'you' must wish to attend graduation. Likewise, there are criteria for (12) and (13). In the two cases, 'you' must plan on attending medical school, must not have applied, and must be worried about taking the MCAT or Organic Chemistry, respectively.

Such criteria for referential candidacy are not unique to *you.* Nunberg (forthcoming) refers to this as the 'granularization of the context'. When we evaluate the potential referents of a referring expression, "we restrict the domain appropriately" (Nunberg forthcoming; also cf. Eggert 1998 for examples with *I, here,* and *now*).

To better understand how *you* is used in public notes, we should consider cases in which 'you' is purely fictional. In Hanks' (1993) research, he asked speakers to explain the meaning of certain deictic terms in Yucatec Mayan. To do so, his consultants invented hypothetical conversations which involved hypothetical 'I's and hypothetical 'you's.

> Although V, DB, and DP animated this speech, its principals are fictional characters, an imaginary 'you' and 'I'. By using first and second person pronouns to create the scene, speakers chose to project themselves and me, their addressee, into the role of interlocutors in the hypothetical framework.
> (Hanks 1993: 136).

We find similar transpositions in novels. First person novels are generally transformations on personal narratives. In personal narratives, speakers narrate events that happened to them to an audience of acquaintances. In first person novels, writers narrate fictional events to an audience of unknowns. Yet embedded within that frame is another frame, in which a fictional narrator is addressing the events to a fictional audience. (Ong (1987) argues that fiction always involves a fictional audience; I argue below that fictional audiences proliferate beyond fiction.) As readers, we are skilled at perceiving the narrative with the eyes of the fictional audience.

In his novella, *Notes from the Underground*, Dostoevsky adds yet another layer of complexity. Rather than transforming a personal narration, he transforms a journal. Journals are an unnatural genre insofar as their authors do not always expect them to be read.

(14) Then - it is still you who are saying this - new economic relations will be established, relations all ready for use and calculated with mathematical exactitude... Of course, it is quite impossible to guarantee (it is I who am speaking now) that even then people will not be bored to tears...
(Fyodor Dostoevsky, *Notes from the Underground*)

Within this fictional journal, the fictional 'I' creates a doubly fictional 'you'. The fictional 'I' explains why:

(15) I, however, am writing for myself, and I should like to make it clear once and for all that if I address myself in my writings to a reader, I'm doing it simply as a matter of form, because I find it much easier to write like that. It is only a form, an empty show, for I know that I shall never have any readers. I have already intimated as much...
(Fyodor Dostoevsky, *Notes from the Underground*)

But what role do we, the actual readers, play in such a (multiply embedded) fiction? At one level we are eavesdroppers, non-ratified recipients of this private journal, but as he creates this doubly fictional audience, Dostoevsky encourages us to inhabit this deeper level addressee role.

(16) Now, of course I might, for instance be taken at my word and asked if I really do not count on any readers, why do I now put down all sorts of conditions, and on paper, too, such as not to pay any attention to order or system, to write down what I remember, etc., etc. Why all these explanations? Why all these apologies? "Ah," I reply, "now you're asking!"
(Fyodor Dostoevsky, *Notes from the Underground*)

Indeed, we are asking. We are asking because we have skillfully traversed the embedded frames and have occupied a fictional addressee role.

We can analyze public notes in a similar fashion. (17) and (18) are successful because the reader cooperates with the writer to create a fictional referent for "you".

(17) Silence is golden.... But without you it's killing us! Strong vocalist wanted for 5-piece alt. rock band We lost our voice in Tennessee Looking for one in SLC
(Posting, Student Union, University of Utah, 11/02)

(18) Stand next to your best friend and read this. One of you could get pregnant before you're ready.
(Advertisement, OSH, University of Utah, 11/02)

In (17), the 'you' the band lacks has certain characteristics that we can infer, characteristics that are absent in the majority of the message's readers. For example, 'you' must have musical talent and aspirations compatible with a five piece alt. rock band. On the other hand, the reader can fill in other details, e.g., the gender and age of 'you'. I expect that most readers create fictional individuals who resemble themselves in all but the pertinent details. Thus I imagined a thirty-something man. Somebody else might imagine a 19 year old woman. (18) is more explicit in the details. Clearly, I am not a good candidate for 'you', since it is physically impossible for me or my best friend to become pregnant before we're ready. Yet I understood the message quite well. The addressee I created is an unattached, heterosexual female undergraduate at the University of Utah, as is her best friend. From her point of view, I then interpreted the rest of the message (which gave advice with respect to birth control etc.). We understand such messages because we inhabit the addressee roles.

It is through this inhabitation that we arrive at a referent for "you". Sometimes, we need do minimal work to arrive at the referent. For example, the vocative in (19) would normally narrow the range of candidates to one.

(19) Abigail, you ate my donut.

I imagine that in many cases we may simply short-circuit the process of inhabiting the addressee role.[3] Nevertheless, the process is an integral part of moving from the sign to its referent.

In other words, I am suggesting a two-stage model for determining the reference of 'you'. This is similar to Nunberg's theory of indexical reference, which involves deferred reference. Nunberg (forthcoming) describes this process as follows:

> We first identify the index of a term—that is, either the demonstratum of a demonstrative or the contextual element that an indexical picks out in virtue of its linguistic meaning—and then proceed to identify the referent of the term, which is something that stands in a salient functional relation to the index.[4]

In trivial cases, "the salient functional relation to the index" is being the index itself. Thus, "you" indexes the person it is addressed to and then may refer to that person. On the other hand, *you* may be used to refer to individuals in a functional relation to the person addressed. Nunberg (1993: 21) gives the following example:

(20) *Chess teacher giving an introductory lesson to a student who has just played 4. N × P...*: According to all the textbooks, you often get in trouble with that move.

In this case, the index is the student, but the referent is chess players in general.

I agree with Nunberg to a point. He is correct that *you* involves deferred reference. However, I suggest that the index is not an entity, but rather a role. In order to determine the referent of a token of *you*, we inhabit the role of the addressee, imagining the characteristics necessary to fit the role. We then evaluate which candidate best fits those characteristics in order to determine the referent.

As I've shown, the actual referent may be purely fictional. In such cases, we compare our real persona to that of the fictional referent we have helped create, evaluating how well they match. If in (16) we were not asking the questions posed by the doubly fictional 'you', then Dostoevsky's fiction would fail. In (7), if we were not moving, we would have no reason to call the author. This becomes even clearer when we consider personal ads.

(21) SUPPOSE YOU'RE MY fantasy? Slender green-eyed blonde headturner who can make you laugh and loves to cuddle seeking best friend, lover, compadre. You're over 5'10", in great shape, a DWM 45-55 who loves his

[3] The term "short-circuit" comes from speech act literature, specifically Morgan (1991).
[4] Nunberg actually goes on to argue that this is not the best way to look at it. Personally, I find the deferred reference model more satisfying than what he replaces it with. Thus in this paper I will continue to use the terminology he employs in Nunberg 1993.

job, the lake, music, movies, theatre, travel. Is this you? Then call this sassy nontraditional attorney.
(Chicago Reader. November 13, 1998. Section 2, p. 46)

In this case, the writer changes footing several times. With the first token of *you*, the referent probably matches most readers. But with the next two tokens the writer increasingly narrows the range of candidates. Finally, she returns to the initial footing, where she asks the 'you' to evaluate his real self with the individual he created for the last token of *you*.

Interestingly, you and I are poor candidates for any of these tokens of *you* because we read the message for scientific, not romantic, reasons. But even when we read scientific writing, we create fictional referents for "you".

(22) Two botanists in conversation would have to establish which lexicon they were drawing on [with respect to fruit]. You and I would be forced to stay with common parlance.
(Clark 1996: 108)

Clark assumes that his readers lack botanical expertise, a reasonable assumption, but not a necessary one. Botanist/linguists or botanist/psychologists could still understand his point, even if they could not be 'you'. Likewise, Dennis Rodman could understand Cecil Adams' point because Rodman can imagine what it would be like not to be Rodman; he can inhabit an addressee role whose referent does not match his identity.

(23) I can accept the idea that you and I made it through [the Millenium/Armegeddon]. But Dennis Rodman?
(Cecil Adams. The Straight Dope. The Chicago Reader, July 3, 1998.)

We need two stages for determining the reference of a token of *you* because the addressee is not a given aspect of the context. All that we can safely presuppose when using *you* is that there is a recipient.[5] The reference is a creative act negotiated by both the sender and the recipient. If it were not creative, we could not misconstrue it; yet (24) and (25) are ubiquitous in our lives.

(24) Are you talking to me?
(25) I wasn't talking to you.

In conclusion, the addressee is a role, not a person. A token of *you* does not refer in itself. Rather, the recipient of a message cooperates in creating a referent. In many cases the referent is the recipient. In other cases, the recipient may deduce that the referent is somebody else (usually another recipient). In examples I've discussed here, the referent is fictional, in which case the recipient must

[5] When "you" falls in a forest and nobody's there, does it have a referent?

evaluate how well their own persona matches that of the fictional referent. We should not regard these fictional uses of *you* as being unusual, though. They are natural extensions of how we always process tokens of the term. To arrive at a referent for "you", a recipient must inhabit the addressee role.

References

Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H., and T. B. Carlson. 1991. Speech acts and hearer's beliefs. In *Pragmatics: A Reader*, edited by S. Davis. Oxford: Oxford University Press.

Clark, H. H., and D. Wilkes-Gibbs. 1990. Referring as a collaborative process. In *Intentions in Communication*, edited by P. R. Cohen, J. Morgan and M. E. Pollack. Cambridge, MA: MIT Press.

Eggert, Randall. 1998. I'm not here now: Saliency in pure indexical reference. In *CLS 34: The Main Session*, edited by M. C. Gruber, D. Higgins, K. S. Olson and T. Wysocki.

Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience*. Cambridge, MA: Harvard University Press.

Goffman, Erving. 1976. Replies and responses. *Language and Society* 5:257-313.

Goffman, E. 1981. Footing. In *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Hanks, William. 1993. Metalanguage and pragmatics of deixis. In *Reflexive Language: Reported Speech and Metapragmatics*, edited by J. A. Lucy. Cambridge: Cambridge University Press.

Hanks, William. 1996. Exorcism and the description of participant roles. In *Natural Histories of Discourse*, edited by M. Silverstein and G. Urban. Chicago: The University of Chicago Press.

Irvine, Judith. 1996. Shadow conversations: The indeterminacy of participant roles. In *Natural Histories of Discourse*, edited by M. Silverstein and G. Urban. Chicago: The University of Chicago Press.

Kaplan, D. 1989. Demonstratives. In *Themes from Kaplan*, edited by J. Almog, J. Perry and H. Wettstein. New York: Oxford University Press.

Kaplan, D. 1991. On the logic of demonstratives. In *Pragmatics: A Reader*, edited by S. Davis. Oxford: Oxford University Press.

Kleinedler, Steve. 2000. Here's what *you* can do: Conversational implicatures triggered by *you*. In *CLS 36: The Main Session*, edited by A. Okrent and J. Boyle.

Levinson, S. C. 1988. Putting linguistics on a proper footing: explorations in Goffman's concepts of participation. In *Erving Goffman: Exploring the Interaction order*, edited by P. Drew and A. Wootton. Boston: Northeastern University Press.

Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

McCawley, J. D. 1984. Speech acts and Goffman's participant roles. ESCOL 1984:260-274.

Morgan, J. L. 1991. Two types of convention in indirect speech acts. In *Pragmatics: A Reader*, edited by S. Davis. Oxford: Oxford University Press.

Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16:1-43.

Nunberg, Geoffrey. Forthcoming. Descriptive Indexicals and Indexical Descriptions. In *Descriptions and Beyond: An Interdisciplinary Collection of Essays on Definite and Indefinite Descriptions and Other Related Phenomena*, edited by M. Reimer and A. Bezuidenhout: Oxford University Press.

Ong, Walter J., S. J. 1987. The writer's audience is always a fiction. In *Twentieth Century Literary Theory : An Introductory Anthology*, edited by D. N. Miller and V. Lambropoulos. Albany: State University of New York Press.

Silverstein, Michael. 1995. Shifters, linguistic categories and cultural description. In *Language, Culture, and Society*, edited by B. G. Blount. Prospect Heights, IL: Wavelantd Press.

Wettstein, H. K. 1991. How to bridge the gap between meaning and reference. In *Pragmatics: A Reader*, edited by S. Davis. Oxford: Oxford University Press.

Shoshoni verb classes and the perception of aspiration*

DIRK ELZINGA
Brigham Young University

1. Introduction.

In Shoshoni, a Numic language of the Uto-Aztecan family, there are a handful of verbal suffixes which show variation in the realization of the initial consonant between a voiceless fricative and a geminate stop; this variation is dependent on the class of the verb to which the suffix is attached and is largely unpredictable synchronically; representative examples are given in (1).

(1) Some alternating suffixes in Shoshoni (Miller 1996)

a. *-kk*^{w}*a/-x*^{w}*a* 'momentaneous'
   [híβikk^{w}a] 'drank up' [tɨ́kkḁx^{w}a] 'ate up'

b. *-kka/-xa* 'resultative'
   [híβikka] 'have drunk' [tɨ́kkḁxa] 'have eaten'

c. *-ppɨnni/-ɸɨnni* 'progressive; frequentive'
   [híβippɨ̀nni] 'be drinking; sip' [tɨ́kkḁɸɨ̀nni] 'be eating; nibble'

The historical explanation for this alternation rests on the interaction of the stress system, degemination, and the optimization of perceptual cues for aspiration. The alternating verbal suffixes in Shoshoni are modern reflexes of historically invariant geminate-initial suffixes. These geminates underwent degemination and preaspiration following short unstressed vowels; this parallels the situation described in Sapir (1930) for Southern Paiute geminate stops. The Shoshoni preaspirated stops underwent a further development and became voiceless fricatives. The conditioning environment for degemination and preaspiration was subsequently lost, leaving behind a lexically determined alternation between geminates and voiceless fricatives (Miller 1980). In this paper I propose that the abductive gesture of the vocal folds which underlies aspiration is more readily perceived on a voiceless fricative than on a preaspirated voiceless stop, and that the change from preaspirated stops to voiceless fricatives in Shoshoni was driven by the perceptual optimization of this gesture.

Support for this analysis comes from a comparison of similar changes in the world's languages; aspirated stops and voiceless fricatives are often related historically in languages around the world; a reasonable explanation for the change from aspirated stop to voiceless fricative is the optimization of the perception of the

* My thanks to John McLaughlin for discussion of this material, as well as to Wick Miller, now deceased, for introducing me to the wonders of Numic phonology. All errors found herein are my own responsibility.

laryngeal abduction gesture; this gesture is transparently recoverable from a voiceless fricative.

The rest of this paper is organized as follows: section 2 provides background in the consonant alternations of Shoshoni and shows the historical origin of the alternating suffixes. Section 3 provides an analysis for the change from preaspirated stop to voiceless fricative and supports this analysis by briefly reviewing similar historical changes in the world's languages. Section 4 is a brief conclusion.

2. Shoshoni final features.

Shoshoni stops undergo certain changes when they stand in phrase-medial position. In initial position they are voiceless and unaspirated; however, in medial position they are variously realized as geminates, as prenasalized voiced stops, or as voiced or voiceless fricatives. This variation depends on the final phonological element of the preceding syllable; these elements are known in Numic studies as the "final features", though it is usual to restrict that term to elements which appear in word-final position. Numic scholars have reconstructed three final features for Proto-Numic: Gemination, Nasalization, and Spirantization.[1] To illustrate, examples of each of these final features are given from Shoshoni; each of the forms in (2) contains a noun followed by the verbalizing suffix *-paʔi* 'have'. The initial *p* of the suffix is realized as a geminate voiceless stop [pp] under Gemination (2a), as a homorganic nasal-stop cluster [mb] under Nasalization (2b), and as a voiced fricative [β] under Spirantization (2c).

(2) Numic final features: Shoshoni[2]
  a. Gemination: [tɨβappaʔi] 'have pine-nuts' (*tɨpa"* 'pine nut', *-paʔi* 'have')
  b. Nasalization: [tsoːmbaʔi] 'have beads' (*tsoːn* 'bead')
  c. Spirantization: [tsoːβaʔi] 'have a great-grandparent' (*tsoː* 'great-grandparent')

Shoshoni (and the other Central Numic languages) have also developed a fourth series, Aspiration, which is historically derived from Gemination, but which is now distinct from it. Under Aspiration, a voiceless stop is realized as a voiceless fricative, often preceded by a voiceless vowel (2d).

(2) d. Aspirated: [haiñt͡ʃi̥ɸaʔi] 'have a friend' (*haintsih* 'friend')

All of the Numic languages show evidence of final features. In Southern Paiute final features regularly appear morpheme-internally, and apply more or less regularly across all morpheme boundaries. In Shoshoni, while morpheme internal occurrences of the consonantal series are fairly regular, they are only preserved across morpheme boundaries for non-verbs. Within the verbal system of Shoshoni the set of final features has been almost completely lost, and all verb stems now end

---

[1] I capitalize the names of the Numic final features in order to distinguish them from more general phonetic and phonological processes which occur in the world's languages.

[2] The final features in Shoshoni are conventionally represented by –" for Gemination, *-n* for Nasalization, and *-h* for Aspiration; Spirantization is not formally represented. The representations for Nasalization and Aspiration presuppose a segmental analysis of the final features; lack of space precludes discussion. See Elzinga (1999) for arguments concerning the segmental representation of the final features in Shoshoni.

with a vowel; the initial consonant of most suffixes surfaces in Spirantized form.[3] There are six verbal suffixes which alternate between a geminate initial consonant and a voiceless fricative initial consonant; these suffixes are shown below in (3) with the verb stems *tɨkka* 'eat' and *hipi* 'drink':

(3) Shoshoni variable suffixes

a. *-kkʷa/-xʷa* 'momentaneous'

[híβikkʷa]	'drank up'	[tɨ́kkḁxʷa]	'ate up'

b. *-kka/-xa* 'resultative'

[híβikka]	'have drunk'	[tɨ́kkḁxa]	'have eaten'

c. *-ppɨnni/-ɸɨnni* 'progressive; frequentive'

[híβippɨ̀nni]	'be drinking; sip'	[tɨ́kkḁɸɨ̀nni]	'be eating; nibble'

d. *-kkandɨ/-xandɨ* 'stative'

[híβikkàndɨ]	'be full (of drink)'	[tɨ́kkḁxàndɨ]	'be full (of food)'

e. *-ttai/-θ̱ai* 'finally'[4]

[híβittài-]	'finally drink'	[tɨ́kkḁθ̱ài-]	'finally eat'

f. *-ttɨɣi/-θ̱ɨɣi* 'start to'

[híβittɨ̀ɣi]	'start to drink'	[tɨ́kkḁθ̱ɨ̀ɣi]	'start to eat'

The choice of suffix depends on the verb stem; some verb stems select the geminate-initial variant, while others select the fricative-initial variant.

All verbs in Shoshoni belong to one of two classes based on the selection of fricative-initial or geminate-initial variants of the alternating suffixes. While there are phonotactic patterns in the verb stems which can help predict a geminate or fricative realization of the suffix-initial consonant, these verb classes are largely opaque to speakers of the language, and are thus lexical in nature. Some examples of verbs of each class are given in (4) below.

(4) Shoshoni verb classes

a. Aspirating

tɨkka	'eat'
na:tia	'race'
tɨpui	'wake up'
wiya:	'dangle'
anni	'fall over'
nanka	'hear'
niikʷi	'say'

b. Geminating

hipi	'drink'
pui	'see'
tsua	'be used up'
mi:	'do'
yɨmi	'swallow'
tɨki	'place'

The alternation between geminated and aspirated suffixes was originally allophonic; these suffixes were underlyingly geminate. Many of these alternating suffixes have cognates in the other Numic languages which are invariably Geminating (5)[5]:

---

[3] There are remnants of a Nasalizing final feature on some verb stems which is evident only before the future suffix *-tuʔi* and the generic aspect suffix *-tɨn*.

[4] [θ̱] is a voiceless non-strident alveolar fricative. Earlier descriptions of Shoshoni referred to this sound as a voiceless tap, but instrumental analysis clearly reveals that it is a fricative.

[5] Kawaiisu distinguishes lenited and non-lenited forms; the non-lenited forms are cognate to Numic geminates.

(5)

Shoshoni	Southern Numic	Western Numic
-ɸɨnni/-ppɨnni		*-ppɨnni* (Northern Paiute)
-xa/-kka	*-ka* (Kawaiisu)	
-x^{w}a/-kk^{w}a	*-k^{w}ee* (Kawaiisu)	*-kku* (Mono)

A regular alternation between geminates and preaspirates is observed in Southern Paiute. In Southern Paiute, a geminate stop degeminates and preaspirates when following a stressless vowel; otherwise, the geminate surfaces as such. A separate process devoices a vowel which occurs before a preaspirated stop (6):

(6) Southern Paiute preaspiration
- a. /pikkák^{w}ittɨ̀a/ 'sore buttocks.ACC'
  pi^{h}káɣ^{w}i^{h}tɨ̀a
  [pi̥káx^{w}i̥tɨ̀ḁ]
- b. /pattákkittùi/ 'cause to burst'
  pa^{h}tákki^{h}tùi
  [pḁtákki̥tùi̥]

In (6a, b), the geminates which close the first and third syllables follow stressless vowels and are thus degeminated and preaspirated. The geminate which closes the second syllable of (6b) follows a stressed syllable and is not subject to degemination.

The same conditions obtained historically for Shoshoni. After the change from geminate to preaspirated stop, the stress shifted removing the complementary environments which gave rise to the alternation. This effectively stranded geminate-initial suffixes in aspirating environments and lexicalized the alternation (7).

(7) Historical Aspiration in Shoshoni

	*tɨ́kka-kk^{w}a	*hipí-kk^{w}a
Aspiration	*tɨ́kka-^{h}k^{w}a	———
Stress Shift	———	*hípi-kk^{w}a

The final development in Shoshoni was from a preaspirated stop to a voiceless fricative (8).

(8) *tɨ́kka-^{h}k^{w}a > [tɨ́kkḁx^{w}a]

As in Southern Paiute, the vowel is devoiced by a separate process. With this as background, I turn now to an analysis of the change from preaspirate to voiceless fricative.

3. Analysis.

In this section I motivate the change from preaspirated stop to voiceless fricative by arguing that perceptual cues for aspiration are more salient on a voiceless fricative than on a stop. The outline of my argument is as follows. First, I assume that aspiration is represented in the grammar by the feature [+spread glottis]. Second, I show that voiceless fricatives are also characterized by the feature [+spread glottis]. Third, I argue that voiceless fricatives are better exponents of the feature [+spread glottis] than aspirated stop are, and that the change from preaspirated stop to voiceless fricative is entirely natural and motivated by optimizing the perception of aspiration.

3.1. Aspiration as [+sg].

I take the feature [+spread glottis] (henceforth [+sg]) to be the grammatical expression of thevocal fold abductive gesture underlying aspiration. In an Optimality Theoretic grammar, preservation of aspiration is expressed by the constraint IDENT$_{IO}$[+sg]:

(9) IDENT$_{IO}$[+sg]: An output correspondent of an input segment bearing [+sg] itself bears [+sg].

I will assume that the promotion of IDENT$_{IO}$[+sg] in the constraint hierarchy accounts for the lexicalization of the Geminate/Aspirate suffix alternation in Shoshoni.

3.2. Voiceless fricatives and [+sg].

In this subsection, I summarize Vaux (1998), which argues that voiceless fricatives bear the feature [+sg]. The evidence for this comes from patterns of assimilation in the New Julfa dialect of Armenian. New Julfa Armenian has a four-way laryngeal constrast among stops (10).

(10) New Julfa consonant system (Vaux 1998: 498)

b	b^h	p	p^h	f	v	m					
d	d^h	t	t^h	s	z	n					
dz	d^{h}z	ts	t^{h}s								
dʒ	d^{h}ʒ	tʃ	t^{h}ʃ	ʃ	ʒ						
g	g^h	k	k^h	χ	ʁ						
				h	ɦ						
							j	ɾ	r	l	ɫ

New Julfa marks the future tense with a prefix *k-* attached to the present subjunctive. This prefix surfaces as [k] before vowels and plain voiceless consonants (11a), as [g] before plain voiced consonants (11b), as [k^h] before voiceless aspirated stops and voiceless fricatives (11c), and as [g^h] before voiced aspirates (11d).

(11) New Julfa future tense (Vaux 1998: 498)

	Underlying form	Surface form	Gloss
a.	k-ert^h-a-m	kert^{h}am	'I will go'
	k-t-a-m	kətam	'I will give'
	k-kien-a-m	kəkienam	'I will exist'
b.	k-bzz-a-m	gəbəzzam	'I will buzz'
	k-l-a-m	gəlam	'I will cry'
	k-zr-a-m	gəzəɾam	'I will bray'
c.	k-t^{h}oʁ-n-ie-m	k^{h}ət^{h}oʁniem	'I will allow'
	k-t^{h}ʃap^h-ie-m	k^{h}ət^{h}ʃap^{h}iem	'I will measure'
	k-χnd-a-m	k^{h}əχəndam	'I will laugh'
	k-savor-ie-m	k^{h}əsavoriem	'I will grow accustomed to'
d.	k-b^{h}ier-ie-m	g^{h}əb^{h}ieriem	'I will carry'
	k-g^h-o-m	g^{h}əg^{h}om	'I will come'
	k-d^h-n-ie-m	g^{h}əd^{h}əniem	'I will put'
	k-d^{h}ziev-ie-m	g^{h}əd^{h}zieviem	'I will form'

From the examples given in (11) it can be seen that the prefix *k-* assimilates in voicing and aspiration to a following consonant. Vaux (1998) interprets this assimilation as evidence for the Laryngeal place node. This node contains at least the features [sg], which is responsible for aspiration, and [voice], which is responsible for voicing. Assimilation of the future tense prefix can now be seen as the spreading of the Laryngeal node of the stem-initial consonant to *k-* (12).

(12) Laryngeal Spreading (Vaux 1998: 499)

k # [+cons]

Laryngeal

If Laryngeal Spreading results in a voiceless aspirated reflex of the future tense prefix before a stem beginning with a voiceless fricative (9c), then the obvious conclusion is that voiceless fricatives bear the feature [+sg] in New Julfa Armenian.

Vaux also provides evidence from synchronic and historical processes at work in Sanskrit that voiceless fricatives bear the feature [+sg]. When a plain voiceless stop is followed by a voiceless fricative, the result is a voiceless aspirated stop/fricative sequence (Vaux 1998: 500-1).

(13)	Underlying form	Surface form	Gloss
	/b^{h}iṣak siːs-ena/	[b^{h}iṣak^{h}siːsena]	healing lead-instrumental
	/ap-su/	[ap^{h}su]	water-locative

The forms in (13) demonstrate that the feature [+sg] spreads from the fricative to the preceding stop in the same manner as in the future prefix assimilation found in New Julfa Armenian.

In the historical development of Pali from Indic (here represented by Sanskrit), fricative/stop sequences are simplified by deletion of the fricative in initial position (14a), and by gemination of the stop in medial position (14b). In each case, the original plain stop is aspirated.

(14)			Sanskrit	Pali	Gloss
	a.	Initial	skand^{h}á-	k^{h}and^{h}a-	shoulder
			stána-	t^{h}ana-	breast
			sparʃa-	p^{h}assa-	touch
	b.	Medial	hásta-	hatt^{h}a-	hand
			yaṣṭí-	yaṭṭ^{h}i-	pole

These changes represent a general simplification of syllable structure that occured between Old Indic (Sanskrit) and Middle Indic (Pali). Sanskrit allows complex onsets and place features in codas, but Pali did not. In the case of onset simplification, an entire segment was lost. In the case of coda deletion, the vacated timing unit was reassociated to the following onset resulting in a geminate. Although the segment was deleted, the featural content of the Laryngeal node was preserved and associated to the remaining consonant. This is a pattern familiar from Autosegmental Phonology (i.e, Autosegmental Stability; Goldsmith 1976: 30-35), and it demonstrates that the feature [+sg] is present on the voiceless fricative.

In light of this evidence, I will assume that the feature [+sg] is universally present on voiceless fricatives and therefore on voiceless fricatives in Shoshoni.

3.3. Optimization of [+sg].

In this subsection I argue that the perception of a [sg] contrast is enhanced if [+sg] is expressed on a voiceless fricative. To this end, I provide cross linguistic data in which aspirated stops have become voiceless fricatives; this suggests that if [+sg] is a gestural implementation of a perceptual [+noise] feature, then a fricative is a better exponent of [+noise] than an aspirated stop (Boersma 1998).

Silverman (1997) contains an extended discussion of the patterns of overlapping and simultaneity of glottal gestures with other gestures such as place of articulation. He observes that languages will stagger or "phase" implementation of glottal gestures with respect to supralaryngeal gestures to optimize their perception. That is, no language will implement a laryngeal abduction gesture to exactly coincide with supralaryngeal closure, as in figure (15). In this figure, laryngeal adduction and the onset of voicing is timed to coincide with the release of the labial stop closure.

(15) Unattested realization of an aspirated *p* (Silverman 1997: 4)

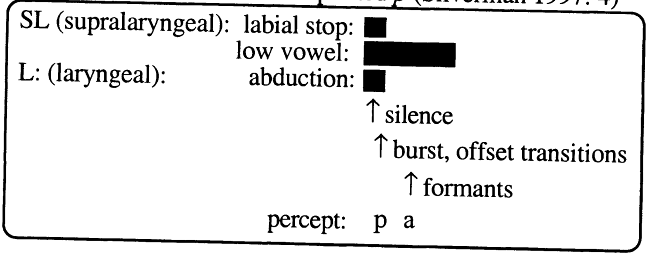
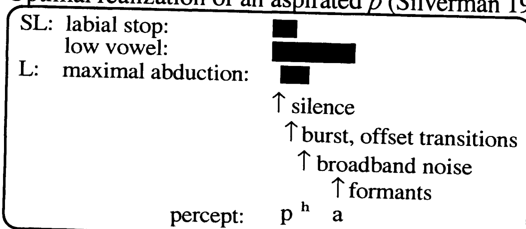


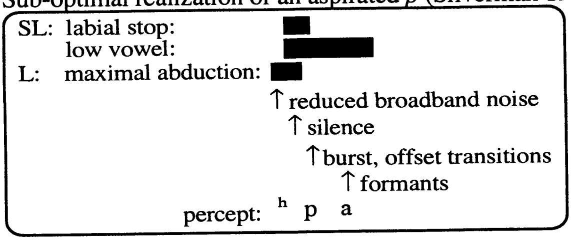
Figure (16) shows a gestural score for an optimally realized aspirated *p*. In this gestural score, the laryngeal gesture significantly overlaps the bilabial closure but also extends beyond it. The onset of voicing thus lags behind the release of the labial closure; this optimizes the perception of laryngeal abduction.

(16) Optimal realization of an aspirated *p* (Silverman 1997: 5)



In (17), a less satisfactory realization of an aspirated *p* is shown. Again, the laryngeal gesture is staggered with respect to bilabial closure, but rather than following bilabial closure, in this case laryngeal abduction precedes it producing a pre-aspirated bilabial stop. This also makes the perception of laryngeal abduction possible.

(17) Sub-optimal realization of an aspirated *p* (Silverman 1997: 6)



Aspirated stops and voiceless fricatives are often related historically in languages around the world. In Classical Greek, there were three series of stops: voiced [b, d, g], voiceless [p, t, k], and voiceless aspirates [p^h, t^h, k^h]. By the end of the fourth century AD, however, the voiceless aspirates had become voiceless fricatives [ɸ, θ, x] (Horrocks 1997: 112-3).

Within the Indo-Iranian branch of Indo-European, where Sanskrit has voiceless aspirates, Avestan has voiceless fricatives (Baldi 1983).

(18) Sanskrit gát^{h}ā-, Avestan gaθao 'song, verse'

In the Pomoan family of languages spoken in California, South Eastern Pomoan shows consistent voiceless fricative reflexes where the other languages have voiceless stops or voiceless aspirated stops (Grekoff 1964).[6]

(19) Proto-Pomoan to Eastern Pomoan

Proto-Pomo	Southern	South Western	South Eastern	gloss
*ʔihp^{h}á	í'pa	—	fa	intestines
*ʔahq^{h}á	á'ka	aká	xa	water
*q^{h}ahbé	ká'be	kabé	xabé	rock
*q^{h}alé	kále	kalé	xalé	tree

In each case, the change proceeds from an aspirated stop to a voiceless fricative. I propose that this change is due to the grounding constraint found in (20) (see Archangeli and Pulleyblank 1994 for a discussion of Grounding and its role in grammar).

(20) SG/CONT: 'If [+sg] then [+cont]; if [+sg] then not [–cont].'

This constraint, along with the constraint IDENT$_{IO}$[+sg] is ranked above the constraint IDENT$_{IO}$[–cont], which exerts pressure on the grammar to preserve an input [–cont] specification. The interaction of these three constraints is shown in the tableau in (21).

[6]My thanks to Mauricio Mixco for bringing this data to my attention.

(21) IDENT$_{IO}$[+sg], SG/CONT » IDENT$_{IO}$[–cont]

	tɨkka-^{h}ka	IDENT$_{IO}$[+sg]	SG/CONT	IDENT$_{IO}$[–cont]
a. ☞	tɨkkaxa			*
b.	tɨkka^{h}ka		*!	
c.	tɨkkaka	*!		

Candidate (21c) fails because an underlying [+sg] is not preserved on the surface. Candidate (21b) fails because the feature [+sg] is realized on a [–cont] segment. Candidate (21a) bests the others since it both preserves the feature [+sg] and realizes it on a segment specified [+cont].

The interaction of constraints in the tableau in (21) shows the role that the realization of [+sg] on voiceless fricatives plays in the historical phonology of Shoshoni. An original preaspirated stop becomes a voiceless fricative in order to optimize the perceptual cues which accompany vocal fold abduction. This change is not unique to Shoshoni but occurs in many of the world's languages.

4. Conclusion.

In this paper I have shown that the historical development of Shoshoni alternating verbal suffixes from preaspirated stop to voiceless fricative has been motivated by considerations governing the perception of aspiration. These considerations dictate that the feature [+sg] is more readily perceived on voiceless fricatives than on voiceless stops. Support for this move has been marshalled from other languages which show a similar historical change. These results provide another instance of the importance of perceptual cues in understanding language change.

The scope of this paper has been modest; but the demonstration of the role of grounding and perception in the change from preaspirated stop to voiceless fricative is important in our understanding of Pre-Shoshoni phonology.

References

Archangeli, Diana and Douglas Pulleyblank. 1994. *Grounded Phonology.* Cambridge, MA: MIT Press.

Baldi, Philip. 1983. *An Introduction to the Indo-European Languages.* Southern Illinois University Press: Carbondale, Illinois.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interaction between Articulatory and Perceptual Drives.* Ph.D. dissertation. University of Amsterdam.

Elzinga, Dirk. 1999. *The Consonants of Gosiute.* Ph.D. dissertation. University of Arizona, Tucson.

Goldsmith, John. 1976. *Autosegmental Phonology.* Ph.D. dissertation. Massachusetts Institute of Technology.

Grekoff, George. 1964. A Note on Comparative Pomo. In Bright, W. ed. *Studies in Californian Linguistics.* University of California Publications in Linguistics 34. Berkeley, CA: University of California Press.

Horrocks, Geoffrey. 1997. *Greek: A History of the Language and its Speakers.* London: Longman.

Miller, Wick. 1980. Preaspirated Consonants in Central Numic. in *American Indian and Indoeuropean Studies: Papers in Honor of Madison S. Beeler.* The Hague:

Mouton.
Miller, Wick. 1996. Sketch of Shoshone, a Uto-Aztecan Language. In *Languages*. ed. Ives Goddard, 693-720. Handbook of North American Indians, vol. 17. William C. Sturtevant, gen. ed. Washington, D.C.: Smithsonian Institution.
Sapir, Edward. 1930. *Southern Paiute, a Shoshonean Language*. Proceedings of the American Academy of Arts and Sciences.
Silverman, Daniel. 1997. *Phasing and Recoverability*. PhD dissertation. University of California, Los Angeles. Published by Garland Press, New York.
Vaux, Bert. 1998. The Laryngeal Specification of Fricatives. *Linguistic Inquiry* **29**: 497-511.

Dirk Elzinga
Department of Linguistics and English Language
Brigham Young University
2079 JKHB
Provo, UT 84602

Dirk_Elzinga@byu.edu

Syntactic and Semantic Valence: Morphosyntactic Evidence from Minangkabau

CATHERINE FORTIN
University of Michigan

0. Introduction

Following Role and Reference Grammar, in this paper it will be claimed that there is more than one type of valence which should be considered in analyzing verbs. There are two distinct, albeit related, types of valence: semantic valence and syntactic valence. I will provide empirical evidence for the theoretical necessity of distinguishing two types of valence with data from Minangkabau, an Austronesian language. A comparison of three passive-like constructions seen in this language supports this distinction. I argue that the distinction between these two types of valence is not only a viable theoretic construct, but is in fact a necessary one, to produce a successful analysis of this system of passive constructions.

1. Valence

The notion of valence (which is sometimes referred to as transitivity) and its implications for argument structure are central to any theory of grammar. However, there has frequently been a lack of clarity surrounding the term, which has traditionally been used to encompass two related, yet discrete, phenomena: co-occurrence restrictions and interpretative phenomena.

In order to illustrate this lack of clarity, consider your own usual interpretation of the term *transitivity*. If you would claim that a verb such as *read* is always transitive, regardless of its syntactic context, you are referring to the verb's *semantic* transitivity, the interpretative phenomenon. Or, if you would claim that *read* is sometimes transitive and sometimes intransitive, depending on its context, you are referring to its *syntactic* transitivity, or its co-occurrence restrictions. Or, to account for both of these, you might instead claim that *read* is listed twice in the lexicon, once as transitive and once as intransitive (Payne 1997: 171).

Several syntactic theories, including Head-Driven Phrase Structure Grammar (HPSG), Lexical Functional Grammar (LFG), and Role and Reference Grammar (RRG), seek to resolve this lack of clarity by clearly distinguishing two distinct types of valence in the grammar itself. The approach followed in this paper is that of RRG, which maintains a difference between semantic valence and syntactic valence. (1) below summarizes the key differences between the two types of valence.

(1) Semantic and Syntactic Valence

Semantic valence	*Syntactic valence*
Determined within the lexicon	Determined by the clause
Thematic structure (an interpretive phenomenon)	Co-occurrence restrictions
The invariant number of semantic arguments that a verb *may* take (Van Valin and La Polla 1997: 147)	The number of arguments present in any given clause, where an *argument* is any nominal element that bears a grammatical relation to the verb (Payne 1997: 170-2)

Croft's (1991: 99) notion of *inherent relationality* is one way of conceptualizing semantic valence, which is perhaps somewhat less tangible than the notion of syntactic valence. As an illustration, consider the following example. Native English speakers have the intuition that a verb such as *read* presupposes the existence of two entities: that who is doing the reading, and that which is being read. Thus, *read* has a semantic valence of two. Conversely, a verb such as *walk* presupposes the existence of just one entity: that which is doing the walking; *walk* has a semantic valence of one. Finally, a noun such as *ball* presupposes the existence of no other entities beside itself. Thus, the semantic valence of *ball* is zero.

In a theory of grammar, these two types of valence should be differentiated because, although they are clearly closely related, they do not always precisely align; whence the lack of clarity surrounding the term as discussed in the previous paragraphs. Speakers frequently encounter cases where the semantic and syntactic valences of a verb are not equivalent. For example, as discussed above, the English verb *read* always has a semantic valence of two. However, its syntactic valence may be one OR two:

(2) Pat read.
(3) Pat read the *Washington Post.*

When an argument is semantically unspecified, as in (2) above, it is not realized syntactically. Certain semantic arguments may be left unspecified, according to the semantics of the verb, where the speaker feels it is permitted by the discourse. This often occurs with verbs which have a sort of 'default' object, as *read.*

There are definite constraints on when, and in which ways, the syntactic and semantic valences are permitted to mismatch. According to the semantics of the verb, semantic underspecification of an argument is not always an option. Compare (4) and (5) below:

(4) Pat wore a red shirt.
(5) * Pat wore.

The precise difference between *wear* and *read* is obviously difficult to quantify, given that both are activity-type verbs with a semantic valence of two. However, it would appear that *wear* does not seem to have a default object in the way that *read* does, thereby causing (5) to be ungrammatical.

Just as it is not always possible for the semantic valence of a verb to exceed its syntactic valence, the opposite restriction holds as well. The syntactic valence of the verb may never exceed its semantic valence, or ungrammatical constructions, such as (7) below, result.

(6) Pat laughed (at the comedy).
(7) * Pat laughed the comedy.

To summarize, there are sentences, such as exemplified in (2) above, in which the semantic valence of a verb is greater than the syntactic valence of the clause in which it appears. In other words, the verb's argument structure, which is lexically specified, is greater than the number of overt noun phrases that are licensed by the verb in the clause in question. These sentences should be able to be clearly accounted for within a syntactic theory.

Further empirical evidence for this claim will be presented in the remainder of this paper. The empirical data is drawn from Minangkabau, an Austronesian language of the Western-Malayo-Polynesian family.

2. Semantic and Syntactic Valence in Minangkabau

Distinct notions of syntactic and semantic valences are reflected in the entire system of Minangkabau verbal affixes (in total, six). Furthermore, native speakers have clear intuitions regarding the morphological distinctions which provide the evidence of this split.

Certain verbal affixes affect *syntactic* valence only. They alter the syntactic distribution of arguments and verbal adjuncts.

Certain verbal affixes further affect the *semantic* valence of the verb. They alter the verb's semantic argument structure.

This distinction is exemplified by a consideration of the different 'passive' constructions in Minangkabau. (For the purposes of this discussion, a 'passive' construction is defined as any construction which reduces the semantic and/or syntactic valence of the verb, thereby allowing 'a "reassignment" of argument NPs from the unmarked configuration of grammatical relations to the desired one' (Croft 1991: 150). These constructions are what will be considered in detail in this paper. In total, there are three prefixes which attach to verbs to generate passive constructions.

One of these prefixes, *di-*, alters only the syntactic structure of the clause.

Two of these prefixes, *ta-* and *bar-,* further alter the semantic structure of the verb; specifically, it may alter the semantic roles that are available to the verb to assign to its arguments.

A preliminary summary of the differences between the various passive constructions is seen in (8) below.

(8) Minangkabau Passives

Affix	*Property*	*Type of Valence Targeted*
di-	Syntactic Passivizer	Syntactic
bar-	'Active' marker	Semantic
ta-	'Involuntary' marker Stative marker	Semantic

With regard to the verbs themselves, the distributions of these morphemes differ (and occasionally overlap), but they never co-occur. In other words, while a given verb may be affixed with, in different situations, *di-*, *ta-,* and *bar-,* there are no cases in which a verb is affixed with more than one of these prefixes simultaneously.

Before turning to the constructions themselves, there is one more theoretical convention to be explained that is assumed in the analysis. This is the notion of semantic macroroles.

2.0. Macroroles

Following Role and Reference Grammar, thematic relations may be distilled down to two generalized semantic roles, or macroroles. The ACTOR macrorole, which represents a generalized agent-type semantic role, subsumes such thematic roles as agent, experiencer, instrument, source and force. The UNDERGOER macrorole, which represents a generalized patient-type semantic role, subsumes such thematic roles as patient, theme, and location (Van Valin and La Polla 1997).

Semantic macroroles are relevant to this analysis of passive constructions in Minangkabau on two levels. First, if we consider an 'active' construction to be one which represents, as per Croft, the 'unmarked configuration of grammatical relations', we may further define the active construction in terms of macroroles. That is, for a semantically transitive verb such as *read,* in the active construction (as seen in (3) above), the 'subject' will bear the ACTOR macrorole and the 'object' the UNDERGOER macrorole. Passive constructions, therefore, represent a deviance of sorts from this unmarked configuration of grammatical relations; generally speaking, the grammatical relations are altered such that the UNDERGOER macrorole is instead borne by the 'subject' of the sentence. Second, the two 'semantic passive' constructions in Minangkabau differ, in part, according to their ability to assign certain macroroles to the verb's arguments.

The semantic passives, which are formed with *ta-* and *bar-,* will be discussed in 2.2. below. First, we will consider the purely syntactic passive, which is formed with *di-*.

2.1. The Syntactic Passive : *di-*

The verbal prefix *di-* is the passivizing morpheme. It reduces the syntactic valence of the verb by one. However, the semantic valence of the verb is not affected. This construction is extremely productive; in discourse, it arguably appears to be preferable to the active counterpart.

Structurally, this construction resembles the passive construction in English. The noun phrase bearing the UNDERGOER macrorole is reassigned to the subject position, preceding the verb. Simultaneously, the noun phrase which bears the ACTOR macrorole is also reassigned to another position in the sentence. If it is a lexical NP, it must be within a prepositional phrase. With a few marked exceptions, the ACTOR macrorole must still be specified within the sentence.

An example of a syntactic passive formed with *di-* is shown below. The active construction, shown in (9), represents the unmarked configuration of grammatical relations: the subject is the ACTOR, the object is the UNDERGOER. Conversely, in the passive construction, shown in (10), the subject is the UNDERGOER, and the ACTOR appears in a prepositional phrase.

(9) anjiang manggigik anak ketek

anjiang	maN[1]- gigik	anak	ketek
dog	TRANS-bite	child	small

'A dog bit the child.'

(10) anak ketek+tu digigik dek anjiang

anak	ketek	+itu	**di-**	gigik	dek	anjiang
child	small	+DEM	PASS-	bite	by	dog

'That child was bitten by a dog.'

Verbs with a semantic valence of less than two (e.g., *run, sleep, arrive*) may not be passivized with *di-*. However, verbs which have a semantic valence of two or more, as the direct result of suffixation of a valence-increasing morpheme, may be passivized with *di-*. An example of this phenomenon is seen below. The verb root, *jago* 'wake', has a semantic valence of one. In the active construction shown in (11), the verb now has a semantic valence of two following the affixation of the causative morpheme *–kan*, which increases the valence by one. The corresponding passive construction is shown in (12). (13) shows that it is ungrammatical to passivize the verb root *jago*, which in its underived form has a semantic valence of one.

Finally, note that in the passive construction in (12), the ACTOR is represented by a full NP in a prepositional phrase; alternatively, the ACTOR may be

[1] The prefix *maN-* marks semantic transitivity.

represented instead by the 3rd person pronominal clitic *nyo*. However, it is not possible for the actor to remain unspecified in a passive (as in the so-called 'agentless passives' seen in English), as demonstrated by the ungrammatical (14).

(11) inyo manjagokan ambo

inyo	maN-	jago	-kan[2]	ambo
3	TRANS-	wake	-CAUS	1

'She woke me up.'

(12) ambo dijagokan dek padusi+tu

ambo	**di-**	jago	-kan	dek	padusi +itu
1	PASS-	wake	-CAUS	by	woman+DEM

'I was woken up by this woman.'

(13) *ambo dijago dek padusitu (14) *ambo dijagokan

In the following section, we shall compare the syntactic passive construction formed with *di-* to the two semantic passives, which are formed with *ta-* and *bar-*. Henceforth, I'll refer to the resulting constructions as *di-passive*, *ta-passive*, and *bar-passive*.

2.2. The 'Semantic Passives': *ta-* and *bar-*

Although these two morphemes, when affixed to verbs, result in passive-like constructions, there are crucial differences between them and the canonical passive-creating morpheme *di-*. These two morphemes further have the ability to alter the macroroles which are assigned by the verb, whereas *di-* only has the ability to alter the configuration in which the macroroles are assigned.

2.2.a. The Involuntary Marker: *ta-*

The first semantic passive is formed with the verbal prefix *ta-*, which has two related functions: it marks lack of volition on the part of the single argument of the verb, and it may further mark the clause as stative. Following the affixation of *ta-*, the verb assigns the UNDERGOER macrorole to its single remaining argument; volition (and in a sense, agency) is thereby removed from the semantic structure. In other words, the argument structure of the verb itself, as it was lexically specified, is somehow altered within a *ta*-passive. Therefore, although this construction appears similar to the syntactic passive formed with *di-*, in that both result in the UNDERGOER macrorole being reassigned to subject position, it is semantically quite different.

There are several differences between the general distribution of *ta-* and the

[2] The suffix *–kan*, here glossed as CAUSATIVE, is very productively used in Minangkabau to add an argument to the verb. In this case, the semantically intransitive verb *jago* 'to wake' is causativized – made transitive – with the affixation of *–kan*.

syntactic passivizer *di-*. Unlike *di-*, *ta-* may be affixed to verbs with a semantic valence of one (as well as those with a semantic valence of greater than one). Furthermore, unlike *di-*, *ta-* may not be affixed to verbs whose valency has been *increased* following derivation, with the suffixation of a valence-increasing morpheme such as *–kan*. Nonetheless, the *ta*-passive is also very productive.

Two examples of *ta*-passives are seen below, the first illustrating a semantically intransitive verb, *lalok* 'sleep', the second illustrating a semantically transitive verb *baka* 'burn'. (15) and (17) represent active constructions containing the verbs; in both cases, as expected, the subject is the ACTOR. The corresponding *ta*-passives are shown in (16) and (18); here, the subject is assigned the UNDERGOER macrorole, as the action is understood to be involuntary. Recall that in (16), the semantic valence of the verb root, *lalok,* is one; following the affixation of *ta-*, the valence is then not reduced, although the semantic role assigned to the single argument undergoes a change from ACTOR to UNDERGOER. In (18), the semantic valence of the verb is reduced by one, from two to one, along with the change in assignment of macrorole to the single argument.

(15) ambo sadang lalok
ambo sadang lalok
1 PROG[3] sleep
'I'm sleeping.'

(16) ambo talalok sajam
ambo **ta-** lalok sa- jam
1 INVOL- sleep one- hour
'I (accidentally) fell asleep an hour.'

(17) si Upiak mambaka sarok
si Upiak maN- baka sarok
NM[4] Upiak TRANS-burn garbage
'Upiak is burning the garbage.'

(18) pohan tabaka
pohan **ta-** baka
tree INVOL- burn
'The trees are burning.'

A further difference between the *di*-passive and the *ta*-passive lies in the constraints on the ways in which the ACTOR macrorole may or may not be specified. In a *di*-passive, the ACTOR generally must be specified[5]; however, in the *ta*-passive, the opposite situation obtains. Only in rare instances is the 'ACTOR' NP to be specified, within in a prepositional phrase. However, in these cases the ACTOR is clearly understood to be a non-agent. Again, this is the key semantic difference between the *ta*-passive and the *di*-passive; there is no volition in the argument structure of the verb with the *ta*-passive. This contrast is shown below with respect to the semantically transitive verb root *bunuah* 'kill'. (19) displays the *di*-passive, (20) the *ta*-passive. In (20) the ACTOR, the first person pronoun *ambo*, is understood to have not killed the chicken on purpose, while in (19), the killing is understood as deliberate.

---

[3] PROG = progressive.

[4] NM = name marker.

[5] There are certain *di*-passives, not discussed here, wherein the ACTOR need not be specified. In order for an agentless *di*-passive to be grammatical, the clause must contain a tense or aspect particle.

(19) kabaw+tu dibunuah dek si Ujan

kabaw	+itu	**di-**	bunuah	dek	si	Ujan
buffalo	+DEM	PASS-	kill	by	NM	Ujan

'That water buffalo was killed (i.e., hunted) by Ujan.'

(20) ayam tabunuah dek ambo

ayam	**ta-**	bunuah	dek	ambo
chicken	INVOL-	kill	by	1

'The chicken was accidentally killed because of something I did.'

The *ta*-passive is further interpreted as representing stativity; that is, it is used to refer to a certain state of affairs. This represents a further contrast between the *di*-passive, which refers to the action itself (not the state of affairs which results from the action) and the *ta*-passive. An example of this is shown below with respect to the semantically transitive verb *buek* 'make'. In (21), the *di*-passive is shown, and in (22), the *ta*-passive is shown; (21) describes the action of making a cake, while (22) describes the state of affairs resulting from the drug being made. It would be ungrammatical to use a *di*-passive to describe a state of affairs such as that depicted in (22).

(21) kue+ko dibuek dek si Ujan

kue	+iko	**di-**	buek	dek	si	Ujan
cake	+DEM	PASS-	make	by	NM	Ujan

'This cake was made by Ujan.'

(22) ubek+tu tabueķ dari akarakaran

ubek	+itu	**ta-**	buek	dari	akaṛ	-REDUP	-an
drug	+DEM	INVOL-	make	from	root	-REDUP	-NOM

'This drug is made from many different kinds of roots.'

2.2.b. The Active Marker: *bar-*

The second semantic passive is formed with the verbal prefix *bar-*, which may be described as the 'active' marker. Pragmatically, *bar-* has the opposite effect of *ta-*, in that it is used when the speaker wishes to emphasize agency on the part of the single argument of the verb[6]. Therefore, in a *bar*-passive, the ACTOR macrorole is assigned to the single argument of the verb, regardless of which role the 'subject' of the verb would be assigned by the verb root. Consequently, if the semantic valence of the verb root is two or greater, the semantic valence will be reduced by one, as the UNDERGOER macrorole is eliminated from the verb's argument structure. As in a *ta*-passive, then, this process alters the argument structure of the

[6] *bar-* is further frequently used to mark certain types of semantically transitive, activity type verbs (in particular, it obligatorily occurs with certain verbs of motion), to indicate iterativity, and to derive certain verbs from nominal roots. These other functions of the affix will not be discussed in this paper.

verb. Finally, *bar-* is used to indicate reciprocity.

The distribution of *bar-* is quite similar to that of *ta-*. Like *ta-*, *bar-* may be affixed to verbs with a semantic valence of one or two. Also like *ta-*, *bar-* may not be affixed to verbs whose valency has been *increased* as a result of derivation. However, this construction appears to have a much more limited productivity than either the *di*-passive or the *ta*-passive.

An example of a *bar*-passive is shown below, with respect to the semantically transitive verb *tapuak* 'clap'. (23) displays the active construction; (24) displays the *bar*-passive, in which the verb has a semantic valence of one, following suppression of the UNDERGOER macrorole.

(23) ambo manapuak meja

ambo	maN- tapuak	meja
1	TRANS-clap	table

'I clapped the table.'

(24) urang urang batapuak

urang	REDUP	**bar-**	tapuak
person	REDUP	ACT-	clap

'The people applauded.'

A *bar*-passive is frequently used when the speaker wishes to imply that the single specified argument is the ACTOR. Therefore, even if another entity is actually responsible for the action described, it would logically never be specified (not even within a prepositional phrase, which is the device utilized by the other two passive constructions). This phenomenon is illustrated below, with respect to the semantically transitive verb *puta* 'spin'. Again, the active construction is shown in (25), and the corresponding *bar*-passive in (26); in (25), the inanimate object *mainan* 'toy' is assigned the UNDERGOER macrorole, while in (26), it receives the ACTOR macrorole.

(25) anak ketek mamuta mainan

anak	ketek	**maN-** puta	main	-an
child	small	TRANS-spin	play	-NOM

'The child spins the toy.'

(26) mainan+tu baputa

main	-an	+itu	**bar-**	puta
play	-NOM	+DEM	ACT-	spin

'This toy is spinning.'

The pragmatic difference between *ta-* and *bar-* is further illustrated below, with respect to semantically intransitive verb *gantuang* 'hang'. In the *ta*-passive in (27), the single argument is the UNDERGOER; the picture, an inanimate object, is hanging involuntarily. In the *bar*-passive in (28), the single argument is the ACTOR; the monkey is understood to be hanging voluntarily.

(27) gambar+tu tagantuang di dindiang

gambar	itu	**ta-**	gantuang	di	dindiang
picture	DEM	INVOL-	hang	on	wall

'This picture is hanging on the wall.'

(28) baruak bagantuang di pohan

baruak	**bar-**	gantuang	di	pohan
monkey	ACT-	hang	on	tree

'The monkey is hanging from the tree.'

A *bar*-passive is also used when the action described is reciprocal. An example of this is shown below, with respect to the semantically transitive verb *paguik* 'hold, hug'. (29) contains the active construction, and (30) the *bar*-passive, which is interpreted as a reciprocal. Semantically, this is similar to English: compare the interpretations of 'Upiak hugged the child' - Upiak is the ACTOR, the child the UNDERGOER - to 'we hugged', where all involved are ACTORS.

(29) si Upiak mamaguik anak ketek

si	Upiak	maN-	paguik	anak	ketek
NM	Upiak	TRANS-	hold	child	small

'Upiak hugged the child.'

(30) kito bapaguik

kito	**bar-**	paguik
2PLINCL	ACT-	hold

'We hugged (each other).'

2.3. Summary of Minangkabau Passives

A summary of the three Minangkabau passive constructions is shown (31) below, along with the relevant aspects of their usage and distribution that were discussed above.

(31) A Summary of Minangkabau 'Passive' Constructions

Affix	*Properties*	*Type of valence targeted*	*Argument structure affected?*	*Occurs with verbs with a minimal semantic valence of*
di-	Syntactic Passivizer	Syntactic	No	Two
bar-	'Active' marker Reciprocal marker	Semantic	Yes	One
ta-	'Involuntary' marker Stative marker	Semantic	Yes	One

3. Conclusion

Empirical evidence of the validity in distinguishing between two types of valence, syntactic and semantic valence, is seen in the array of passive-type constructions exhibited by Minangkabau. In total, there are three types of passive-like structures.

The first passive construction, the *di*-passive, behaves like a canonical passive construction. Only the syntactic valence of the verb is reduced, as a reconfiguration of grammatical relations is all that is involved. The UNDERGOER macrorole, which is normally assigned to the object of the verb, is reassigned to its subject; simultaneously, the ACTOR macrorole is assigned to a noun phrase which must be licensed by a prepositional phrase, as the verb no longer has the ability to do so. The argument structure of the verb, as it is lexically specified, is not itself affected.

The other two constructions, the *ta*-passive and the *bar*-passive (the so-called 'semantic' passives), further modify the semantic structure of the verb. The argument structure of the verb itself, as it is lexically specified, is affected in these two constructions. In each case, the semantic valence of the verb is reduced, not merely its syntactic valence. In the first semantic passive, in which the verb is affixed with *ta-*, the INVOLUNTARY marker, the UNDERGOER macrorole is assigned to the single argument of the verb. In the second semantic passive, in which the verb is affixed with *bar-*, the active marker, the ACTOR macrorole is instead assigned to the single argument of the verb. The macroroles assigned within the *ta*-passive and *bar*-passive are independent of the macroroles assigned by the verb itself outside of these constructions.

In this paper, it is claimed that this system of passive constructions can be successfully analyzed only in light of the distinction between semantic and syntactic valence as proposed by HPSG, LFG, and especially RRG. As such, it may be suggested that such a theoretical distinction has the potential to provide a straightforward solution for other long-standing problems concerning argument structure. Furthermore, this distinction is supported by the richly differentiated system of morphosyntax seen in Minangkabau.

References

Croft, William. (1991). *Syntactic categories and grammatical relations: the cognitive organization of information.* Chicago: The University of Chicago Press.

Payne, Thomas E. (1997). *Describing morphosyntax: a guide for field linguists.* Cambridge: Cambridge University Press.

Van Valin, Robert. and La Polla, Randy. (1997). *Syntax: structure, meaning and function.* Cambridge: Cambridge University Press.

Catherine Fortin
Department of Linguistics
4080 Frieze Building
105 South State Street
University of Michigan
Ann Arbor, MI 48109-1285

fortinc@umich.edu

A study of the speaking fundamental frequency characteristics and perceived pitch characteristics of Black and White women

NICOLE GASKINS
University of California, Los Angeles

1. Introduction

Since the 1980's a modest body of literature normalizing and comparing the speaking fundamental frequency (SFF) characteristics of Black speakers has been amassed. However, this literature has not reached a consensus on whether or not the intonation of Black and White speakers differs. Several researchers have reported that Black speakers tend to use a lower SFF, although not all of these studies concluded that those differences were statistically significant (Hudson 1977, Holbrook and Hudson 1981, Holbrook and Hudson 1982, Wheat and Hudson 1988, Awan and Mueller 1996, Xue and Mueller 1996, Sapienza 1997, Hagstrom et al. 2001). Other studies report the exact opposite, finding that Black speakers have a significantly higher SFF (Tarone 1973, Walker 1981). Some studies report no differences in the SFF level of Black and White speakers (Mayo and Manning 1994, Grant and Mayo 1995, Xue and Mueller 1996, Morris 1997, Hagstrom et al. 2001). A number of studies report greater variability in Black intonation (Tarone 1973, Hudson 1977, Hudson and Holbrook 1981, Hudson and Holbrook 1982, Morris 1997). Other studies refute this claim (Mayo and Manning 1994, Awan and Mueller 1996, Hagstrom et al. 2001).

It is even more difficult to form an adequate description of the SFF characteristics of Black speakers who are female from the even sparser and more contradictory literature that specifically measures the SFF of adult Black women. The author of this study has discovered only nine such studies, four of which are for elderly populations. Hudson (1977) found that Black women have lower mean SFF levels and higher SFF ranges and standard deviations than White women. Holbrook and Hudson (1982) also found that Black women have a lower SFF than White women and that White women have a smaller SFF range and standard deviation. The reading fundamental frequencies of Black women were also found to be lower than those of White women and the reading frequency range of Black women was also found to be higher than that of White women (Hudson and Holbrook 1981). The Black women in Tarone's (1973) study, however, used a higher pitch level, while their pitch ranges were greater than White women's. Sapienza (1997) reported no differences in the SFF of Black and White women. The majority of studies on elderly Black women report that they use lower SFF than elderly White women (Dejarnette and Holland 1993, Xue and Mueller 1996, Hagstrom et al. 2001) but there was one report of higher SFF among elderly Black women (Walker et al. 1981). Several problems may be the cause preventing the firm formation of a description of the SFF characteristics of Black women: these studies vary in instrumentation, data analysis, types of speech analyzed, and methods of comparison, factors which should alone lead to discrepancies in results.

The current body of literature on intonation leads to the conclusion that American intonation needs be reexamined as a function of race and that more data needs to be collected for Black women. The purpose of this study is to 1) collect SFF measures for samples of both Black and White women, 2) describe the SFF characteristics of Black women, 3) determine if actual differences exist in the SFF characteristics of Black and White women by comparing these measures, and 4) to determine whether any perceived differences exist in the pitch characteristics of Black and White Women.

2. Hypothesis

Intonation has been a subject of interest in the field of gender linguistics since its beginnings, as an awareness of the acoustical differences in the speech of men and women existed even in folk linguistics. Researchers of sex differences in language state three main intonational characteristics of female speakers of American English; 1) women use greater ranges of intonation, 2) women use higher SFF levels, and 3) women prefer variable intonation patterns (McConnell-Ginet 1983). While these patterns may hold true for some groups of society, there has been no compelling evidence to show that these patterns are universal, as is suggested by the language of previous works. Like researchers of American intonation, investigators of gender linguistics have largely been content to overlook racial differences and study racially skewed samples. Consequently, one must question if conclusions about the intonation of female speakers can be generalized to accurately describe the speech of all American women.

Given the influence of society on the structure and use of language, it seems very likely that even if similar patterns of intonation were found in all American female speech communities, among certain social groups there would exist consequential variability. Intonation has proven to be sensitive to several social variables, including gender (Brend 1972) and social status (Gregory et al. 2001), and so why should it lack sensitivity to such social constructs as race? My hypothesis is that as the result of their distinct social histories and realities there is a difference in the SFF characteristics of Black and White American women; it is hypothesized that Black women in upwardly mobile social positions use lower and less varied intonation than their White counterparts, and that they, like White males, imitate White female pitch characteristics in a derogatory manner.

3. Method

3.1. Subjects and procedures: Part I

This study was conducted in two parts. In Part I of the study a sample of twenty-three women were self-selected from a population of employees at the L.A. County Department of Children and Family Services. Twelve of the subjects were White women and eleven were Black women (racial classification was dependent upon answers reported in a demographic survey). The subjects ranged in age from 24 to 45 (so as to exclude those women who might be affected by adolescent or menopausal voice change), with a median age of 35 for the White women and a median age of 39 for the Black women. All of the subjects were native speakers of English and had been residents of Southern California for at least ten years.

Recordings of these subjects' speech samples took place in a quiet office. Subjects were instructed to give a job description and an account of their typical workday, speaking as naturally as possible and for at least two minutes. Several subjects had trouble speaking continuously for two minutes, preventing the researcher from collecting a full two minutes of continuous speech in six samples. The quality of two recordings may have been affected by participants speaking too low or by recordings being taken at an improper volume. Following the recordings subjects were administered demographic and mood questionnaires.

3.2. Subjects and procedures: Part II

In the second part of the study a sample of nine women were taken from a population of female undergraduate students at UCLA. All of the subjects were friends of the researcher. All of the subjects were Black and all were non-smokers. All of the subjects had been residents of Southern California for at least two years; five of the subjects having been natives of Southern California, three of the subjects were from Northern California, and only one subject was from out of state. The subjects ranged in age from 19 to 22 with a median age of 20.

Recordings of speech samples for the second part of the study took place in the subjects' homes or in the homes of their friends. Subjects were first instructed to silently read a passage. Then subjects were instructed to read the same passage out loud, pretending as if they were "Keesha", a hypothetical woman. Subjects were then instructed to read the same passage out loud once again, but this time pretending as if they were "Becky", a hypothetical white woman.

3.3. Equipment and data analysis

Recordings were made with a Maranatz professional cassette recording deck (Model PMD501) and a Shure microphone (Model SM58). Eighteen second samples were taken form the midsection of the recordings from part one of the experiment. Samples from part two of the study included the entire passage and varied in length from 14 to 17.5 seconds. PC Quierer version 6.1 (PC based speech analysis software developed and distributed by Sciconrd) was used to obtain pitch analysis graphs from which measures of mean pitch were manually calculated. The mean SFF, SFF standard deviation, and SFF range for each speech sample was calculated. Then these numbers were averaged to determine mean values for each group.

4. Results

4.1. Part I

The mean average SFF was 156.7 Hz with a standard deviation of 26.31 Hz for Black subjects. The mean average SFF was 170.3 Hz with a standard deviation of 22.03 Hz for White subjects. The mean upper limit for Black subjects was 201.9 Hz and 225 Hz for White women. The mean lower limit was 123.8 Hz for Black subjects and 135 Hz for White subjects. A 2 sample T tests at 5% showed the difference in mean SFF between Black and White subjects is not statistically significant ($t = 1.31$, $p = .20$).

The mean SFF for the Black subjects of this study is considerably lower than that previously reported. Hudson and Holbrook (1982) reported that the mean modal SFF for Black women is 188.85 Hz. This was the only study to measure spontaneous speech for Black women whose ages (18-29 years) overlapped with the ages of the subjects of the present study (24-45 years). The difference in the measures obtained in this study and in Hudson and Holbrook's study may be accounted for by age differences of subjects and the use of mean versus modal values. The mean SFF for White subjects is considerably lower than those values previously reported. Stoicheff (1981) reported that the mean SFF for White women 20-29 years old was 224.3 Hz, 213.3 Hz for women ages 30-39, and 220.8 for women ages 40-49. Snidecor (1951) and Linke (1953) reported that the mean SFF's of adult white females are 213.5 Hz and 199.8 Hz, respectively. These studies measured subjects reading a standard passage, while the measures in the current study involve spontaneous speech, possibly accounting for the considerable discrepancy in average values. Also, Stoicheff controlled for the smoking variable, while smoking was not controlled for in the present study.

The mean SFF standard deviation was 14.8 Hz for Black subjects and 19.49 Hz for White subjects. A 2 sample T tests at 5% showed that the difference in mean SFF standard deviation between Black and White subjects is not statistically significant ($t = 1.66$, $p = .11$).

Three Black subjects and four White subjects were cigarette smokers. Since one of the effects of cigarette smoking on speech has been found to be a lowering of SFF (Sorensen 1981), the seven cigarette smokers in the current study were under suspicion of having altered SFF levels. However, a comparison of the mean SFF of smokers versus nonsmokers reveals no evidence that the smokers in the present study produced different SFF levels than the nonsmokers in the present study.

Despite efforts to obtain a uniform population of university educated professional workers, one White subject and eight Black subjects were not social workers. These women were clerical workers whose education level ranged from high school to two-year College. A comparison of the mean SFF of the Black clerical (153 Hz) and social workers (166.45 Hz) revealed no statistically significant differences between the two groups (t = -.62, p = .58).

4.2. Part II

The two speech samples (one as a Black woman and one as a White woman) for each of the subjects were directly compared. All nine subjects used an appreciably lower SFF level when speaking as a Black woman as opposed to when they imitated a White woman, increasing their SFF level by an average of 35 Hz when imitating a White woman. For each of the nine subjects the increase was statistically significant (t = -2.6, p = .02). The mean SFF for the speech productions as Black women was 231.1 Hz with a standard deviation of 25.9 Hz. This value is much higher than that reported by Hudson and Holbrook (1982) and can probably be accounted for by the theatrical nature of the speech task in this study. The mean SFF for the imitations of White women was 266.3 Hz with a standard deviation of 30.8 Hz. This value is also greater than the average values previously reported for White women but it lies within the range of mean fundamental frequencies previously reported for white women (Fitch and Holbrook 1970, Stoicheff 1981).

The mean SFF standard deviation for subjects speaking as Black women was 55.8 Hz. The mean SFF standard deviation for the imitations of White women was 60.2 Hz. This difference is not statistically significant (t = -.56, p = .58). In fact, seven out of the nine subjects altered the SFF standard deviation between their Black and White imitations by only a negligible degree. The other two subjects demonstrated a significant increase in SFF standard deviation when imitating a White woman.

5. Discussion of results

5.1. Comparison of Part I and Part II results

The SFF and SFF standard deviations obtained for the Black subjects in Part I of the present study are significantly lower than the SFF and SFF standard deviations obtained for subjects in Part II of the present study. This may be accounted for by three factors: age differences (Stoicheff (1981) showed that SFF decreases from early adulthood to middle age), differences in speaking tasks (SFF has been shown to vary with speech activity (Grant and Mayo 1995:32)), and differences in recording settings (intonation may vary with social situation (Tarone 1973:36)).

5.2. Part I

The results of this study indicate that Black and White women do not differ significantly in SFF levels. Instead the majority of mean SFF's for both Black and White women lie within the same overlapping range (130 to 190 Hz).

There was a tendency for the Black subjects to speak at lower SFF levels than White subjects. The Black subjects had a lower mean upper limit and a lower mean lower limit when compared to White subjects. Five of the Black subjects had a mean SFF under 150 Hz, while only two white women spoke under 150Hz.

Seven of White subjects had mean SFF over 170 Hz, while only four Black women spoke in that same range. Emphasizing this distribution pattern where the mean SFF's of the Black subjects gathered toward lower frequencies and the mean SFF's of the White subjects are gathered toward higher frequencies were outlying mean SFF's. The Black subjects account for the outliers at the low end. Conversely, the White subjects account for the outliers at the high end.

The results also reveal that Black and White women do not differ significantly in SFF variability (measured by standard deviation). Still, the Black subjects did have noticeably lower mean SFF standard deviations than did the White subjects. Only two Black subjects had a SFF standard deviation above 20 Hz, while six White subjects did. Three Black subjects had SFF standard deviations below 10 Hz, while only one White subject spoke with such little frequency variability. The SFF standard deviations of subjects followed a distribution pattern similar to the one previously mentioned, where Black subjects tended to gather toward the higher frequencies.

It is possible that statistically significant differences were not found in the mean SFF values of the Black and White women sampled in the present study because the Black subjects were not of a uniform social class. The inclusion of a large number of clerks in the sample of Black women may have attenuated any racial effects on SFF characteristics. It would appear that the results of the present study invalidate this claim, as statistical analysis revealed that there is no difference in the SFF characteristics of clerical and social workers, and Black social workers had some of the higher mean SFF's. However, the sample of Black women was relatively small, and by dividing this sample into even smaller groups for comparison, the validity of any comparisons made is compromised.

5.3. Part II

The results from Part II of the present study show that Black women do indeed perceive White women as speaking at higher pitch levels than they do themselves. When the Black subjects imitated White women, they increased their SFF levels significantly.

The results also indicate that Black women perceive White women as speaking with similar or increased pitch variability (measured by SFF standard deviation). The majority of subjects kept their SFF variability constant between Black and White imitations, although some subjects increased SFF variability greatly when imitating White women. When the speech samples were grouped by race, the mean SFF standard deviation was slightly higher for the imitations of White women than for the imitations of Black women.

5.3. Part I and II in tandem

The imitative behaviors of the subjects indicate that there is a perception of pitch characteristics that may not wholly reflect actual speaking behavior. Black subjects significantly increase their SFF level and in some cases their SFF variability when imitating White women, despite the fact that most Black and White women display similar SFF characteristics. The sharp contrast that Black women expressed between the imagined SFF characteristics of White women and their own parallels the dispersion patterns of extreme mean SFF's. Black women comprised the outliers and the majority of speakers at lower SFF levels and SFF standard deviations while White women comprised the outliers and the majority of speakers at higher SFF levels and SFF standard deviations.

I propose that these subjects are conveying greater differences in the pitch characteristics of Black and White women than actually exist because they perceive a conspicuous difference in the pitch characteristics of Black and White women. I further suggest that this altered perception is due to a focus on the extreme speakers and that it is the marginal and not the majority that exerts influence over perception when Black women hear the speech of White women.

6. Analysis and implications

6.1. Racial effects on intonation

Consistent with the findings for the elderly (Xue and Mueller 1996, Hagstrom et al. 2001), children (Awan and Mueller 1996, Wheat and Hudson 1988), and young adults (Hudson 1977, Holbrook and Hudson 1982), the Black women in this study tended to speak with lower fundamental frequencies than White women. Consistent with the findings of Xue and Mueller (1996) and Hagstrom et al. (2001), Black women have lower SFF standard deviations than White women. The results of this study do not show a statistically significant difference the SFF characteristics of Black and White, but the dispersion patterns of SFF values point to the conclusion that race does matter in intonation. Several previous studies have failed to find a difference in the SFF characteristics of Black and White speakers (Mayo 1990, Mayo and Manning 1994, Grant and Mayo 1995), but these studies were not designed in a way to investigate language in its social context. For the socially triggered effects that I hypothesized to be displayed in speech, speakers would have to be involved in social interaction such as the interview or role-play tasks used in the present study. Studies that obtain measures from reading or sustained vowel production are not suitable for detection of what are, I suggest, socially-conditioned racial effects on SFF.

6.2. The speech of the powerless

The results of this study indicate that there is a slight tendency for professional Black women to speak with lower and less varied SFF than their White counterparts. These racial differences can easily be explained using the same theories of the relation between power, language, and society used to explain sexual differences intonation. For the intonational patterns established for White women imply instability, a lack of self-control, and inferiority (Cameron 1982:74, McConnell-Ginet 1982), and Black women are a historically oppressed and socially disadvantaged group who are fighting to avoid such negative stereotyping. A rejection by professional Black women of this “powerless” language is an attempt of a "powerless" group in society to assert power, command respect, and effect upward social mobility. Common sentiments in this community are that "if you are a Black woman on the job, you're probably fighting marginalization, invisibility and often out-and-out racism from nine to five" and that in the competitive business world White women have "the power... of blonde hair and blue eyes" (Golden 2002:190,194). These beliefs that the same resources available to White women are not available to themselves have given professional Black women the incentive to seek out language as an alternative means of power acquisition.

Not only has it been shown that voice fundamental frequency has a function in communicating information about social status (Gregory et al., 2001), but lowering SFF level and variability does appear to provide certain social rewards to women, namely greater credibility and access to authority and professional mobility (Kramer 1974, Thorne and Henley 1975:19). In fact, several magazines, assertiveness training courses, and self-help books advise professional women to lower their pitch in order that they might be taken more seriously (Cameron 1992). The results of this study combined with previous knowledge about women's intonation imply that Black women have taken heed of the negative stereotyping of female speech and are capitalizing upon the power strategy of avoiding White women's intonation in order to advance in the professional world.

One may ask why then are White women content to continue using the intonation labeled instable, inferior, and lacking in self-control? They are probably not; previous studies show that lowering intonation is a power strategy that may also be employed by White women. This may well be implied by the fact that the SFF values for professional White women in this study were considerably lower than the values for nonprofessional White women in previous studies. The results of this study, however, indicate that Black women emphasize their use of certain SFF characteristics even more than White women. I suggest that this is because Black women feel that they must work twice as hard to achieve the same. Further experiments controlled for professional status are necessary for both races to verify this point.

The results of the second part of this study revealed that Black women speak with higher SFF levels and sometimes-greater SFF variability when imitating White women. The fact that Black women change their SFF characteristics when imitating White women demonstrates that Black women perceive a difference in the pitch characteristics of Black and White women and have stereotyped the speech of Black and White women. As mentioned earlier, high and variable pitch has several negative associations. This is especially true in the Black English speech community, where high pitch is associated with dishonest discourse and low pitch is associated with true discourse (Morgan 1996:415). I suggest that Black women emphasize the above mentioned pitch characteristics in the speech of their White counterparts while deemphasizing them in their own speech because they connect negative associations with White women and positive associations with themselves. When White men imitate White women by producing high and varied pitch it is done in a derogatory fashion (McConnell-Ginet 1982:74). The results of the present study indicate when Black women emulate the pitch characteristics of White women it also has negative implications. Several of the subjects adopted a mocking attitude when imitating a White woman, indicating that to speak like a White woman is viewed as derogative and that their high and varied intonation patterns are viewed negatively.

6.3. Stereotypes

The discrepancy between differences in SFF characteristics of Black and White women and the differences in pitch characteristics perceived by Black women begs us to ponder the question: How are stereotypes created? As the results of this study illuminates, the explanation of the birth of stereotypes is tied to the very nature of perception.

While no statistically significant differences were found in the SFF of Black and White women, there were differences in the dispersion patterns of their SFF. There was a tendency for Black women to speak at lower SFF levels and to lay as outliers on the low end and for White women to speak at higher SFF levels and to lay as outliers on the high end. There was also a tendency for Black women to have lower SFF standard deviations and to lay as outliers on the low end and for White women to have higher SFF standard deviations and to lay as outliers on the high end. Perhaps it is the distribution of extreme speakers that act as a catalyst in the formation of the stereotype that White women speak at higher pitches and with greater pitch variability. As Black women take notice of the pitch of only the speakers who display extreme pitch characteristics and as the qualities associated with these characteristics coincide with Black women's beliefs, perception pushes Black women to perceive White women as speaking with high and variable intonation and Black women as speaking with low and level intonation despite the fact that most Black and White women speak with similar SFF characteristics.

Adler et al. (1983) have proposed an analysis of the process of perception. In our everyday tasks of organizing the world, we are bombarded with far more stimuli than we have the capacity to process. Since we cannot focus on every single sensation in the world, the first steps of perception involves a process of selecting which stimuli to pay attention to and which to ignore. In order to manage the plethora of stimuli, the human brain focuses on stimuli that stand out, in this case, speakers with extreme intonation patterns. Another factor influencing our selection of which stimuli to focus on is our motives (Adler et al. 1983: 56). It can be speculated that Black women are motivated to perceive White woman as having high and variable (or " powerless") pitch because of their desires to attain power amid racial competition. Black women perceive White women (even those White women who speak with a normal SFF level and normal SFF variability) as speaking extremely because they do not notice the majority of speakers who display normal pitch characteristics, instead, driven by motives, they focus on those who stand out and generalize about the whole population based on a few, thus leading to the creation of a stereotype.

7. Conclusion

The results of this study have led to two major conclusions:
1) While there are no statistically significant differences in the SFF characteristics of Black and White women, Black women had lower mean mean SFF levels, and lower mean SFF standard deviations. In addition, the summary statistics for each racial group tended to be dispersed around opposite ends of the continuum. These results have greater validity than those of previous studies because the current study uses spontaneous speech samples as opposed to reading samples or prolonged vowel production and because the comparisons of the current study were made using data collected at the same time and in the same manner for both racial groups. 2) There are perceived differences in the pitch characteristics of Black and White women. Black women perceive White women as speaking at a much higher pitch level and in some cases with much greater pitch variability than they do themselves.
Three explanations are suggested for:
1) Black women have devalued the pitch characteristics traditionally associated with "women's speech". 2) Professional Black women avoid these speech patterns as a means of self-assertion and power acquisition. 3) The stereotypes that Black women have formed about White women's pitch characteristics are due to the very nature of perception, whereby attention is focused on the conspicuous and influenced by motives.

Studies of speaker race identification (Lavner et al. 2001) and cases of housing discrimination (Lundy and Massey 2001) demonstrate that suprasegmental differences between the races are perceptible and have significant societal implications. This study further supports the notion that pitch characteristics are perceptually salient and are instrumental in social interaction. It is imperative that language continues to be studied as a function of race, both for the sake of a more comprehensive understanding of its usage and because it leads us to insight about more general issues pertaining to race and society.

References

Adler, Ronald B, Lawrence B. Rosenfeld and Neil Towne. 1983. Interplay: The process of interpersonal communication. New York: CBS College Publishing.

Awan, Shaheen N. and Peter B. Mueller. 1996. Speaking fundamental frequency characteristics of White, African American, and Hispanic kindergartners. *Journal of Speech and Hearing Research* 39: 573-577.

Brend, Ruth M. 1972. Male-female intonation patterns in American English. In Henley, Nancy; Thorne, Barrie (eds.) *Language and sex: difference and dominance*, 84-88. Rowley, MA: Newbury House.

Cameron, Deborah. 1992. Feminism and linguistic theory. St Martin's Press: New York.

DeJarnette, Glenda and Richard W. Holland. 1993. Voice and voice disorders. In D Battle (eds.) *Communication disorders in mulitícultural populations*, 212-238. Boston: Andover Medical Publishers.

Fitch, James L and Anthony Holbrook. 1970. Modal vocal fundamental frequency of young adults. *Archives of Otolaryng* 92: 379-382.

Golden, Marita. 2002. White Women At Work. Essence. 33. 190-198.

Grant, William C. and Robert Mayo. 1995. Fundamental frequency, perturbation, and vocal tract resonance characteristics of African-American and White American males. *Journal of the National Black Association of Speech-Language-Hearing* 17: 32-38.

Gregory, Stanford W, Brian E. Green, Robert M. Carrothers, Kelly A. Dagan and Stephen W. Webster. 2001. Verifying the primacy of voice fundamental frequency in social status accommodation. *Language & Communication* 21: 37-60.

Hagstrom, Fran, Jianping Hao, Richard Neeley ans Steve Xue. 2001. Speaking fundamental frequency characteristics of elderly Euro-American and African-American speaker: building a clinical comparative platform. *Clinical Linguistics & Phonetics* 15: 245-252.

Holbrook, Anthony and Amelia I. Hudson. 1981. A study of the speaking fundamental frequency of young Black adults. *Journal of Speech and Hearing Research* 24: 197-201.

Holbrook, Anthony and Amelia I. Hudson. 1982. Fundamental frequency characteristics of young Black adults: spontaneous speaking and oral reading. *Journal of Speech and Hearing Research* 25: 25-85.

Hudson, Amelia I. 1977. A study of the speaking and reading fundamental vocal frequency of young Black adults. Tallahassee, Florida: Florida State University dissertation.

Kramer, Cheris. 1974. Women's speech: separate but unequal?. *Quarterly Journal of Speech* 60: 14-24.

Lavner, Yizhar, Judith Rosenhouse and Isak Gath. 2001. The prototype model in speaker identification by human listeners. *International Journal of Speech Technology* 4: 63-74.

Linke, Eugene. 1953. A study of pitch characteristics of female voices and their relationship to vocal effectiveness. Ames, Iowa: SU of Iowa dissertation.

Lundy, Garvey and Douglas S. Massey. 2001. Use of Black English and racial discrimination in urban housing markets. *Urban Affairs Review* 36: 452-469.

Mayo, Grant. 1990.Vocal fundamental frequency and vowel formant frequency charcteristics of normal African-American and White-American adults. Memphis, Tennessee: Memphis State University dissertation.

Mayo, Robert and Walter H. Manning. 1994. Vocal tract characteristics of African-American and European-American adult males. *Texas Journal of Audiology and Speech Pathology* 20: 33-36.

McConnell-Ginet, Sally. 1983. Intonation in a man's world. In Henley, Nancy; Kramarae, Cheris; Thorne, Barrie (eds.) *Language, gender, and society*, 69-88. Cambridge: Newbury House Publishers.

Morgan, Marycliena. 1996. Conversational Signifying. In Ochs, Elinor; Schegloff, Emanuel A; Thompson, Sandra A (eds.) *Interaction and grammar*, 405-434. England: Cambridge UP.

Morris, Richards J. 1997. Speaking fundamental frequency characteristics of 8 through 10-year-old White and African-American boys. *Journal of Communication Disorders* 30: 101-116.

Sapienza, Christine M. 1997. Aerodynamic and acoustic characteristics of the adult African American voice. *Journal of Voice* 11: 410-416.

Snidecor, John C. 1951. The pitch and duration characteristics of superior female speakers during oral reading. *Journal of Speech Hearing Disorders* 16: 44-52.

Sorensen, David. 1981. The effects of long term smoking on selected voice characteristics. *Dissertation Abstracts International*. 41: 3324.

Stoicheff, Margaret L. 1981. Speaking fundamental frequency characteristics of nonsmoking female adults. *Journal of Speech and Hearing Research* 24: 437-441.

Tarone, Elaine E. 1973. Aspects of intonation in Black English. *American Speech* 48: 29-36.

Thorne, Barrie and Nancy Henley (eds.) 1975. Language and sex: difference and dominance. Rowley, MA: Newbury House.

Walker, Virginia G. 1981. Speech and language characteristics of an aging population. In Lass, Norman J. (eds.) *Speech and language: advances in basic research and practice*, 143-202. New York: Academic Press.

Wheat, Marcia C. and Amelia Hudson. 1988. Spontaneous speaking fundamental frequency of 6-year-old Black children. *Journal of Speech and Hearing Research* 31: 723-725.

Xue, An and Peter B. Mueller. 1996. Speaking fundamental frequency of elderly African-American nursing home residents: preliminary data. *Clinical Linguistics & Phonetics* 10: 65-70.

Nicole Gaskins
3120 Bagley Ave #3
Los Angeles, CA 90034

ngaskins@ucla.edu

The Status of the Morpheme in Georgian Verbal Morphology[1]

OLGA GUREVICH
University of California, Berkeley

1 Introduction

There are at least three traditional, widespread approaches to inflectional morphology: item and arrangement (IA) (Lieber 1992), item and process (IP) (Steele 2002), and stem and paradigm (SP) (Stump 2001, Anderson 1992). In addition, some models combine features of more than one type (such as Distributed Morphology, Halle and Marantz 1994). The differences between the three approaches have to do with representing individual words, parts of words (morphemes), and patterns of inflections (paradigms). All of these approaches have tried to analyze Georgian as an example of a complex morphological system; thus, Georgian has been used as a kind of benchmark for theories of morphology. The goal of this paper is twofold: to compare existing approaches and to provide a satisfactory analysis of Georgian verbal morphology.

The Georgian morphological system is complex and relatively unusual typologically. As will be shown below, morpheme-based accounts of it have been forced to make arbitrary decisions in order to account for the particular combinations of morphemes in each word form. Ideally, however, a morphological description should not have to make unmotivated decisions. Instead, I propose a word-based approach where the smallest meaningful unit is a word, not necessarily a morpheme. This approach, I will argue, resolves the analytical challenges presented by Georgian, and is in line with current psycholinguistic evidence.

I will use the Word-and-Paradigm (WP) approach (Matthews 1991, Blevins 2003) which is based on pedagogical grammars and operates with whole words and example paradigms. This meta-theory is necessarily an idealization of storage and pattern extraction mechanisms employed by actual speakers; thus, the proposed morphological analysis is meant to be maximally compatible with what is known about language use and acquisition, but not to represent the mental

[1] I would like to thank Jim Blevins, Andreas Kathol, and Alice Harris for their detailed, insightful comments. I am grateful to Shorena Kurtsikidze and Vakhtang Chikovani for sharing their language with me. All errors are my sole responsibility.

lexicon of any particular speaker. In what follows, matters of principle will be distinguished from matters of convenience[2].

The main difficulty that Georgian presents for traditional approaches is as follows. Any attempt to assign separate meanings to individual morphs leads to very abstract or very ambiguous meanings. In a sense, the meaning of a word is not straightforwardly composed from the meanings of its constituent parts. The meaning of the form determines the particular combination of parts, in contrast with other possible combinations. Thus, knowing what one wants to say guides the selection of the morphs, without necessarily attributing meaning to those morphs. This general property is manifested in several ways, to be exemplified below. Traditional approaches which assume meaningful morphemes have been forced to make arbitrary decisions in various parts of the paradigms. Instead, an analysis that takes words as minimal units can do a much better job. In the words of Matthews (1991:204, speaking of the much more familiar Greek and Latin),

> Many linguists tend to boggle at such systems. They seem complicated, while agglutinating systems seem so simple. They may even seem perverse. Why should a language have rules which obscure the identity and function of its minimal elements?
>
> An apologist for ancient grammar would answer that these elements are fictions. They are created by the modern method; and, if we foist them on a flectional system, we are bound to describe it as an agglutinating system that has somehow gone wrong. In the ancient model the primary insight is not that words can be split into roots and formatives, but that they can be located in paradigms. They are not composed of simple parts, but are themselves the parts within a complex whole. In that way, we discover different kinds of relation and, perhaps, a different kind of simplicity.

This paper argues for a top-down, rather than bottom-up, approach, where the larger constructions determine the combination of morphs (form elements) for each word, and where the meaning of each word is not necessarily directly composed from the meanings of the parts[3]. Entire words are stored, organized into example paradigms. Each paradigm cell represents an abstraction; it is a mini-construction which combines semantic properties (tense, agreement) with form properties (a combination of stem and affixes). Morphological generalizations lie in the strength of example paradigms and are reinforced by the frequency of individual verbs and verb types. No abstract rules are necessary in this model, apart from the generalizations extracted on the basis of individual examples. It is important here to make the distinction between units of storage (words) and units of analysis (words or morphs), or between storage and representation for the purposes of linguistic analysis. This kind of analysis avoids stipulations to resolve slot competition or to determine the order of rule application.

[2] The model should be general enough to be applicable to languages other than Georgian, as well. Nothing in the suggested mechanisms of storage and pattern extraction is language-specific, but rather relies on what is known about acquisition in general.

[3] This approach is in line with a view of syntax espoused by Construction Grammar (Kay and Fillmore 1999, Goldberg 1995).

The rest of the paper provides an overview of the Georgian verbal system and examines the problems that it causes for traditional approaches. The striking feature of the verb is that larger constructions (such as tense) determine the combination of subparts (morphs) for a given form, but most individual morphs do not determine larger constructions. In what follows, I will describe several sets of such morphs: the pronominal agreement markers (Section 3), preverbs (Section 4), and versionizers (Section 5). For each set, I examine the difficulties that arise for a morphemic analysis, and suggest a word-based alternative that avoids those difficulties. Section 6 concludes and provides some further discussion.

2 The organization of the Georgian verb

The Georgian verb can reflect (agree with) the subject, the object, and in some cases the indirect object or the recipient (beneficiary). The verb can also be marked for other parameters, including causative, tense and aspect. The agreement markers marking various arguments fit into slots before or after the stem, and they compete for these slots. Thus the verb can be morphologically very complex.

To further complicate the matter, the same set of affixes can mark agreement with the syntactic subject or object, depending on the verb and the tense of the verb. Nominal arguments (subject, object, and indirect object) of a verb show up in different cases depending on the grammatical class and the tense of the verb. Thus, the syntactic properties of a verb can also play a role in its morphology, and the relation between syntactic and morphological properties is not always straightforward.

In general, the structure of a Georgian verb appears to follow the template in (1). Not all of the parts of the template show up in all forms of a verb, but when they do show up, the order is always the same. As mentioned above, one of the main difficulties will be assigning meanings to individual morphemes; however, for expository convenience morphemic glosses will be used where possible.

(1) preverb+AGR1+theme Vowel+stem+thematic suffix + tense/aspect + AGR2
(Cf. Hewitt 1995)

A preverb marks contrasts in aspect and tense. There are several preverbs, and they are lexically associated with the verbs; a semantic connection is no longer transparent. The theme vowel is used to indicate the relationship between subject and indirect object in some situations, but is "high-jacked" to mark other contrasts elsewhere. The thematic suffix, like the preverb, is lexically associated with the verb, and for the most part marks inflectional class.

3 Pronominal agreement markers

Traditionally, agreement markers in Georgian are divided into several sets. For the sake of convenience, I will use the standard terminology. Each set includes values for slots AGR1 and AGR2. For the uninverted verbs, the Set A markers indicate subject agreement, and Set B markers indicate object agreement. These sets will make it easier to talk about cases where the association between verbal agreement and syntactic arguments is different.

Set A:

	Singular	Plural
1	v-	v- -t
2	-	- -t
3	-s/-a/-o	-an/en/es/nen[4]

Set B:

	Singular	Plural
1	m-	gv-
2	g-	g- -t
3	-/h/s[5]	-/h/s -t[6]

3.1 Slot competition in the pre-stem position

I will first look at regular agreement marking for Transitive verbs. An example of a present-tense paradigm follows.

(2) Forms of the verb "to draw" in the present tense:

	OBJECT				
Subj	**1SG**	**1PL**	**2SG**	**2PL**	**3**
1SG	--	--	g-xat'av	g-xat'av-**t**	v-xat'av
1PL	--	--	g-xat'av-**t**	g-xat'av-**t**	v-xat'av-**t**
2SG	**m**-xat'av	**gv**-xat'av	--	--	xat'av
2PL	**m**-xat'av-**t**	**gv**-xat'av-**t**	--	--	xat'av-**t**
3SG	**m**-xat'av-**s**	**gv**-xat'av-**s**	g-xat'av-**s**	g-xat'av-**t**	xat'av-**s**
3PL	**m**-xat'av-**en**	**gv**-xat'av-**en**	g-xat'av-**en**	g-xat'av-**en**	xat'av-**en**

From the pre-stem elements in chart in (2), it can be seen that *m-* indicates 1sgObj agreement, *gv-* indicates 1PlObj agreement, and *g-* indicates 2Sg and 2PlObj agreement. On the other hand, *v-* indicates 1Sg or 1PlSubj agreement. Thus, the AGR1 slot does not seem to correspond to a particular morphosyntactic property or argument. It sometimes expresses properties of the subject, sometimes of the object. In a theory that aims to derive forms based on morphosyntactic properties, how does one organize the rules in order to get the correct derivation? The generalization here seems to be that the slot AGR1 expresses object properties for 1st and 2nd person objects, and subject properties otherwise. A theory that aims to make the rules "fall out" right would be forced to stipulate that the object properties for certain kinds of objects supersede the subject properties. That is, in fact, what one sees in the literature.

---

4 The alternation between different 3rd person subject markers is conditioned phonologically in some cases, lexically and syntactically in others (i.e. different variants are chosen for different tenses).

5 The alternation between *h-* and *s-* is phonologically conditioned. The alternation between the zero-allomorph and h/s is a bit more complicated, and I will not discuss it here.

6 Not all verb types mark number distinctions for 3rd person objects, i.e. 3Sg and 3Pl obj forms are often the same.

Anderson (1992) organizes his rules in such a way that rules realizing direct object agreement apply before those realizing subject agreement. This arrangement of rules and rule blocks produces the correct results; however, the decision of which rules are placed in which blocks is not motivated by external (synchronic) factors.

Similarly, Stump (2001) resolves the problem by making the realization rules for agreement apply in a particular order. Rather than assigning rules to rule blocks, however, rules can apply in expanded or unexpanded form, and the expanded rules (being more specific than non-expanded rules) apply first. Thus, the *g*- prefixation rule applies in expanded form, but the *v*- prefixation rule applies in non-expanded form. This distinction results in the correct forms being produced; however, there does not seem to be any motivation for this distinction other than to get the correct forms.

3.2 Slot competition in the post-stem position

The post-stem slot AGR2 often indicates subject properties. Looking now at the post-stem elements in table (2), one notices the following. Starting from the bottom row, the marker *-en* seems to indicate 3pl subjects; the marker *-s* seems to indicate 3sg subjects in all cases but one. The marker *-t* consistently marks 1Pl and 2Pl subjects. However, the same marker *-t* that marked plurality of 1st or 2nd person subject markers, also sometime shows up with 1st or 2nd person *singular* subjects. Whenever that happens, the object has to be plural. In the paradigm in (2), three forms are exactly the same, yet they mark four different combinations of subject and object properties (1PlSubj/2SgObj, 1PlSubj/2PlObj, 1SgSubj/2PlObj, 3SgSubj/2PlObj). This homophony is not specific to a particular verb, but rather is systematic in the language.[7]

The problem, again, is that the AGR2 slot is not associated with a particular morphosyntactic property or argument. The morph *-t* is especially problematic. One possibility is that it simply marks plurality of any argument, subject or object. However, there are forms with a plural argument that do not have *-t*: for example, the 2SgSubj, 1PlObj *gv-xat'av* has a plural object. But the form *gv-xat'av-t* already exists; it means 2PlSubj, 1PlObj; so adding *-t* onto the former would make it identical to the latter. On the other hand, the 2SgSubj, 1PlObj form is already distinguished from its 1SgObj counterpart *m-xat'av* by a different prefix. Thus, in a sense, the appearance of *-t* marks contrast, and not simply plurality. This is where morphemic analyses are unsatisfactory. Halle and Marantz (1994) postulate a *-t* insertion rule wherever there is a plural argument, and a *-t* deletion rule for the circumstances where *-t* is expected but does not appear. Stump (2001) posits three different morphemes *-t*; however, there is no historical or synchronic evidence for positing different morphemes. Neither approach seems

[7] For some verbs, a stem alternation occurs in 3rd person subject forms, reducing the homophony to only three forms. However, the AGR1 and AGR2 slots are still filled the same way for those verbs.

very plausible as an actual strategy employed by a language user, as they are forced to make arbitrary decisions.

The problems with slots AGR1 and AGR2 are only problems if we try to assign individual meaning to each slot or each morph that appears in those slots. If, however, we operate with entire words and compare related word forms, it does not matter what each individual part of the word means, as long as the forms are different. This approach does not need to make arbitrary decisions regarding the number of *t*'s.

Carmack (1997) suggests that instead of associating morphemes with morphosyntactic properties, a morphosyntactic property is associated with a *verb frame* (a kind of circumfix, the combination of AGR1 and AGR2 slots). E.g. 1Pl Subj/3 Obj is marked by the verb frame *v____t*, 3Pl Subj /1Sg Obj is marked by the verb frame *m____en*, etc. I borrow his terminology here. However, it should be emphasized that the frames have no ontological status but are only abstractions from entire stored words; that is, I am *not* arguing for agreement as a set of circumfixes instead of prefixes and suffixes.

Each cell in a paradigm can be seen as a construction which combines meaning properties (such as 1SG Subject, 2PL Object, Present tense) with a set of formal properties (AGR1 morph, AGR2 morph, stem, etc.) In that sense, the "frames" in Carmack's sense are simply formal parts of these constructions, which also include other morphs. The view of cells as mini-constructions permits non-compositional combinations of morphs, but allows some morphs to be compositionally transparent, as well. In this analysis, slot competition is no longer an issue because for each cell in the paradigm, it is already known which parts are to be used. This may seem unsatisfactory because it does not *explain* why the morph *g-* overrides the morph *v-*. While that is a very important question and should be addressed, I believe that such explanations must be diachronic (see Harris 1985) and are not part of the synchronic grammar that speakers possess.

3.3 System of verbal tenses (screeves) and Inversion

So far we have looked only at the present tense. The Georgian verbal tenses (called *screeves*, from the Georgian *mc'k'rivi* meaning 'row') are organized into three series based on the syntactic realization of the nominal arguments (case marking) and the morphological realization of the verb forms. Not all verbs have forms in all three series, and the syntactic realization of nominal arguments in each series depends on the particular verb class. However, for all verbs all screeves within a series have the same syntactic arguments and the same agreement patterns.

3.3.1 Series I:

The First Series is historically the imperfective series. Case marking in this series is Nom/Dat for all verbs except verbs with lexical Inversion (see Section 3.3.3).

Example:

(3) k'ac-i dzaɣ'l-s xat'- av- s
man-NOM dog-DAT draw-T.S.-AGR2
"The man draws / is drawing the dog." (Present)

3.3.2 Series II:

The Second Series, historically perfective, includes the Aorist Indicative and the Aorist Subjunctive screeves.

The case marking in this Series in the modern language appears to be based on the active/stative distinction: active subjects are marked by the ergative case, while non-active subjects are marked by the Nominative case (see Harris 1985). Example (PV stands for preverb):

(4) k'ac-ma dzaɣ'l-i da- xat'- a
man-ERG dog-NOM PV-draw-AGR2
"The man drew the dog." (Aorist Indicative)

Comparing verb forms in Series I and Series II, there does not seem to be a specific morpheme that signals a difference in tense[8]. The preverb marks perfectivity and also appears in Future tense (which is part of Series I). The morphs in AGR1 are exactly the same for both Series. The only difference is the absence of the thematic suffix in Series II and different AGR2 morphs for 3rd person subjects. However, the thematic suffix also disappears in certain Series III forms (see next section), and the morphs in AGR2 mark verbal agreement - i.e. there isn't one morpheme that appears in all of the Series II forms. So it seems that the only way one can tell that these forms belong to Series II is by contrasting them with forms in Series I: the presence vs. absence of the thematic suffix in conjunction with the choice of an AGR2 morph. This comparison between whole words is a very natural operation for language learners, and provides further evidence for a word-based approach.

The AGR1 morphs are the same in Series I and II. The distribution of *-t* in AGR2 is also the same. However, the AGR2 morphs associated with 3sg and 3pl subjects are different. Thus the meaning of those morphs has to combine the tense/aspect and morphosyntactic properties of the subject. Alternatively, the "verb frames" that involve 3rd person subjects are less general than frames for other verb forms: they are specific to tense/aspect and are thus less semantically transparent.

The distribution of AGR1 and AGR2 markers in the Series II is in (5).

[8] This problem has gone largely unnoticed in previous analyses.

(5) Forms of the verb "to draw" in the Aorist Indicative tense:

	OBJECT				
Subj	**1SG**	**1PL**	**2SG**	**2PL**	**3**
1SG	--	--	da-**g**-xat'-**e**	da-**g**-xat'-**e**-**t**	da-**v**-xat'-**e**
1PL	--	--	da-**g**-xat'-**e**-**t**	da-**g**-xat'-**e**-**t**	da-**v**-xat'-**e**-**t**
2SG	da-**m**-xat'-**e**	da-**gv**-xat'-**e**	--	--	da-xat'-**e**
2PL	da-**m**-xat'-**e**-**t**	da-**gv**-xat'-**e**-**t**	--	--	da-xat'-**e**-**t**
3SG	da-**m**-xat'-**a**	da-**gv**-xat'-**a**	da-**g**-xat'-**a**	da-**g**-xat'-**a**-**t**	da-xat'-**a**
3PL	da-**m**-xat'-**es**	da-**gv**-xat'-**es**	da-**g**-xat'-**es**	da-**g**-xat'-**es**	da-xat'-**es**

Another issue for traditional approaches is whether to analyze the post-stem vowel in Series II as part of the AGR2 slot or as part of the stem. If it is part of the stem, then separate stem alternations would have to be learned for each Series, as well as for different subject properties. If it is part of the AGR2 slot, then it appears that in cases where *-t* shows up, the slot contains two morphemes, not just one. In the analysis proposed here, the vowel is a part of the word form and is learned along with the rest of the form. It does not need a separate ontological status, and thus it does not matter whether or not it is considered part of the stem. Rather, the constructions where this vowel appears license it as part of the form.

3.3.3 Series III: Inversion

The Third Series includes the Perfect (also known as the First Evidential or First Resultant) and the Pluperfect (also known as the Second Evidential or Second Resultant). The main difference between this series and the other two is the 'evidential' or 'apparent' connotation: the speaker was not present during the time of the event, and thus takes no responsibility for what is being said. In Series III, the subject and object agreement markers seem to be reversed, as in (6).[9]

(6) Forms of the verb 'to draw in the 1st resultant tense (Series III)

	SUBJECT		
OBJ	**1SG**	**1PL**	**2SG**
1SG	--	--	da-**g**-i-xat'av-var
1PL	--	--	da-**g**-i-xat'av-var-**t**
2SG	da-**m**-i-xat'av-xar	da-**gv**-i-xat'av-xar	--
2PL	da-**m**-i-xat'av-xar-**t**	da-**gv**-i-xat'av-xar-**t**	--
3	da-**m**-i-xat'av-s	da-**gv**-i-xat'av-s	da-**g**-i-xat'av-s

[9] Notice that the table is organized differently: subject properties appear in columns, and object properties appear in rows.

OBJ	SUBJECT		
	2PL	**3SG**	**3PL**
1SG	da-**g**-i-xat'av-var-**t**	da-**v**-u-xat'av-var	
1PL	da-g-i-xat'av-var-**t**	da-**v**-u-xat'av-var-**t**	
2SG	--	da-u-xat'av-xar	
2PL	--	da-u-xat'av-xar-**t**	
3	da-**g**-i-xat'av-**t**	da-u-xat'av-s	da-u-xat'av-**t**

In Series I and II, it was possible to assign specific meanings to the AGR1 morphs *m-*, *g-*, *gv-*, and *v-*. However, as Series III shows, that meaning does not hold in all cases. The morphs *m-*, *g-*, and *gv-* mark properties of the object in Series I and II, but properties of the subject in Series III for transitive verbs. Conversely, the morph *v-* marks properties of the subject in Series I and II but properties of the object in Series III. Again, constructional paradigm cells (verb frames) which combine subject / object properties with the tense / aspect properties are needed. The Inversion in Series III is not random: it corresponds to a change in the syntactic valence of a verb. The Inversion construction is similar to a passive, although the exact change in the semantic roles played by the subject and the object is unclear.

(7) k'ac-s dzaγl-i t'urme da-u- xat'- av- s
man-DAT dog-NOM apparently PV-T.V-draw-T.S.-AGR2
"The man has apparently drawn the dog." (1Res)

The association between agreement markers and person/number properties remains constant in Inversion, but the mapping onto syntactic argument structure is reversed. Thus, Inversion is essentially a problem of linking semantic roles (such as Agent, Patient, Perceiver, and Stimulus) to syntactic arguments (Subject and Object) and can be resolved if one adopts a level of argument structure which is separate from the level of semantic roles. This split has been adopted in HPSG, LFG, and Relational Grammar, and is not a problem for a morphological description. For the Georgian case, any association between agreement markers and person/number properties can be deterministic, but the linking to syntactic arguments is dependent on the series and/or verb type.

4 Preverbs

As mentioned above, certain distinctions in the Georgian tense system are marked by preverbs. Most verbs have a preverb lexically associated with them, although there is also a group of verbs that do not have preverbs. In Series I, tenses without preverbs are Present Indicative, Past Continuous, and 1st Subjunctive. Adding preverbs to these forms produces forms of Future, Past Frequentative (used in conditional constructions), and Future Subjunctive, respectively. The preverbed forms systematically differ from their preverb-less counterparts in aspect (Holisky 1981). However, aspectual differences alone do not account for the differences in tense. Examples in (8) demonstrate the present and future screeves which differ by a preverb.

(8) a) surat-s **xat'av-s**
picture-DAT draw.3sgSubj.Pres
'He draws / is drawing a picture' (*Present*)

b) surat-s **da-xat'av-s**
picture-DAT draw.3sgSubj.Fut
'He will draw a picture' (*Future*)

The preverbs are also used in all Series II and all Series III tenses. Most of these forms do not have preverb-less equivalents.

The particular tense construction determines whether or not a preverb will be used, and the choice of preverb is lexically determined. A constructional view of tenses avoids the problem of having to assign meaning to the preverbs, or expecting preverb meanings to account for tense meanings: the presence of a preverb is determined by the tense construction, but the meaning of the preverb does not need to be combined with the meaning of the rest of the form to get the meaning of the whole.

5 Versionizers

Indirect object (or beneficiary) agreement is realized in Georgian by the so-called "version vowels" which appear between the AGR1 slot and the stem. For most transitive verbs where a beneficiary argument is semantically plausible, the version vowel is used to mark agreement in Series I and II. The vowel *-i-* indicates that the beneficiary is 1st or 2nd person, i.e. a conversation participant, or identical with a 3rd person subject (so-called "subjective version") (9), *-a-* is neutral version (10), and *-u-* marks that the beneficiary is a 3rd person different from the subject ("objective version") (11).

(9) saxl-s v-i-šeneb
house-DAT build.1sgSubj.Pres
'*I* build / am building a house *for myself.*'

(10) saxl-s v-a-šeneb
house-DAT build.1sgSubj.Pres
'*I* build / am building a house (beneficiary not specified).'

(11) saxl-s amxanag-s v-u-šeneb
house-DAT friend-DAT build.1sgSubj.Pres
'*I* build / am building a house *for a friend.*'

However, in Series III, the version vowel no longer marks IO agreement. Instead, the vowel in that slot becomes a screeve marker: the vowel *-i-* appears in 1st and 2nd person subject forms of the 1st Resultant, *-u-* appears in 3rd person forms of 1st Resultant, and the vowel *-e-* consistently appears in all forms of the 2nd Resultant. The beneficiary in these tenses appears in a postposition phrase. In the following 1st Resultant form, the thematic vowel *-i-* appears where the beneficiary is 3rd person.

(12) saxl-i a-m-**i**-šeneb-ia
house-NOM build.1sgSubj.1Res
'I have (apparently) built a house (beneficiary not specified).'

For this class of verbs, the versionizer is the only morph which distinguishes 1st Resultant forms from corresponding 2nd Resultant forms. Without getting into the detailed analysis of the versionizers, the key point for this paper is that the versionizer can only be given meaning within the context of a particular tense. Thus, again, the bigger construction (tense) determines which smaller parts (morphs) will be used.

6 Conclusion and Further Remarks

From a preliminary examination of the Georgian data, it seems clear that a morphemic approach is not ideal. It is very difficult to assign meaning to individual morphemes, and where such assignment is possible, the decisions that have to be made are in many cases arbitrary.

Analyzing entire words seems to be a better approach. Entire "verb frames" can be associated with combinations of agreement and tense/aspect properties. Such an approach eliminates the need to assign meaning to individual morphemes. It also avoids stipulations regarding slot competition. The idea of verb frames also circumvents the problem of splitting words into morphemes and assigning parts of the verb to agreement slots of parts of the stem. Thus what I am suggesting is not an alternative segmentation of the forms, but a very different way of looking at the issues, in which segmentation is not a primary concern.

In general, systems in which the meaning of a word is not strictly composed of the meaning of its constituent morphemes are not uncommon. In such systems, the combination of the morphs determines the meaning of a form because it stands in contrast to other possible combinations of morphs. A word-based model would be well-suited to represent such systems. Such a model should be made compatible with evidence from language acquisition and results of other psycholinguistic testing.

References

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge; New York: Cambridge University Press

Blevins, James P. 2003. "Inflectional classes and economy." To appear in L. Gunke, G. Müller and G. Zinfonun (eds.) *Explorations in Nominal Inflection.* Berlin: Mouton de Gruyter

Carmack, Stanford. 1997. "Blocking in Georgian Verb Morphology." In *Language*, 73 (2): 314-338

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: The University of Chicago Press.

Halle, Morris, and Alec Marantz. 1994. "Distributed Morphology and the Pieces of Inflection." In K. Hale and S.J. Keyser (eds.), *The View from Building 20.* Cambridge: The MIT Press

Harris, Alice C. 1985. *Diachronic Syntax: The Kartvelian Case*. Orlando: Academic Press.

Hewitt, B.G. 1995. *Georgian: A Structural Reference Grammar.* Amsterdam/Philadelphia: John Benjamins

Holisky, Dee Ann. 1981. "Aspect Theory and Georgian Aspect." In P. Tedeschi and A. Zaenen (eds.) *Syntax and Semantics, V. 14: Tense and Aspect*. Newy York, London: Academic Press; 127-144

Kay, Paul and Fillmore, Charles. 1999. The *what's x doing y* construction. *Language* 75 (1): 1-33.

Lieber, Rochelle. 1992. *Deconstructing Morphology: Word Formation in Syntactic Theory.* Chicago: University of Chicago Press

Matthews, P.H. 1991. *Morphology*. Cambridge: Cambridge University Press

Steele, Susan. 2002. "Many Plurals: Inflection, Informational Additivity, and Morphological Processes" in P. Boucher (ed.). *Many Morphologies.* Cascadilla. Somerville, MA; 82-108.

Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press

Olga Gurevich
1203 Dwinelle Hall
University of California, Berkeley
Berkeley, CA 94720-2650

olya@socrates.berkeley.edu

Definition of the base[1]

S. L. ANYA HOGOBOOM
University of California, Santa Cruz

1. Introduction

McCarthy and Prince define the base for morphological affixation in a two part definition, given in (1). I refer to this as the "single-side definition of the base."

(1) McCarthy and Prince 1993a:106

> (a) The base of a suffixed morpheme is the phonological string preceding the exponent of that morpheme, up to the nearest initial edge [of a PrWd.
> (b) The base of prefixed morpheme is the phonological string following the exponent of that morpheme, up to the nearest final edge] of a PrWd.

This definition was meant to apply both to cases of fixed morphological affixation and to cases of reduplicative affixation. However, Generalized Alignment (McCarthy and Prince 1993b) made it unnecessary to refer to a base for fixed morphological affixation (non-reduplicative affixation). For example, in McCarthy and Prince 1993a, the placement of the Ulwa possessive affix (examples are given in (2)) was accounted for in part through the notion that the base of the possessive is a foot.

(2) Ulwa Possessive Forms (Hale and Lacayo Blancho 1989)

siwanak	'root'	siwa-ka-nak	'his/her root'
kululuk	'woodpecker'	kulu-ka-nak	'his/her woodpecker'
bas	'hair'	bas-ka	'his/her hair'

In McCarthy and Prince 1993b this same data is accounted for through an alignment constraint ALIGN (*ka*, L; FT, R), which serves to align the left edge of

[1] I would like to thank Judith Aissen, Junko Itô, Afton Lewis, Jaye Padgett, Caro Struijke, and Florence Woo as well as the participants of the 2002 research seminar at UCSC, and the audience members of the UCSC mini-conference and of BLS 29. Any mistakes are of course my own.

the affix to the right edge of a foot. The base no longer needs to be referred to in order to account for fixed segment affixation.

The definition of base in (1) is still employed, however, to account for the placement of reduplicative morphemes (e.g Kager 1999). In this paper, I discuss some of the difficulties with this definition. I propose an alternative definition of the base and show that it does not run into the same difficulties as the single-side definition. I argue for a definition of the base such as that in (3).

(3) Base (proposed definition): The base of a reduplicative morpheme consists of all of the segments in the output with the exception of the reduplicant.

2. Argument for a uniform base

I will argue that the proposed definition of the base (as given in (3)) is preferable to the more commonly assumed, single-side definition (as given in (1)), based on three main arguments.

First, the proposed definition does not handle prefixes differently from suffixes. In the most straightforward cases, like the examples in (4), the definition in (3) and that in (1) pick out the same base. Therefore the more general one should be preferred.

(4)	Agta:	tak	'leg'	tak-takki	'legs'	(Healey 1960)
	Dakota:	haska	'be tall'	haska-ska	'be tall, pl.'	(Shaw 1976)

Second, it is not clear that the single-side definition can pick out the correct base in all cases. For example, Nahuatl (Sullivan 1988, Chisholm 2001) has a pattern of reduplication (5).

(5) Nahuatl Diminutive (Plural)

(a)	(i)	siwa-pil woman-DIM	'dear little woman'
	(ii)	siwa-pi̱-pil woman-PL-DIM	'dear little women'
(b)	(i)	tutu-pil bird-DIM	'little bird'
	(ii)	tutu-pi̱-pil bird-PL-DIM	'little birds'

In the plural forms, the root is followed by the reduplicant, which is followed by the diminutive suffix. Since the reduplicant is a suffix (it follows the root) the single-side definition picks out the string of segments preceding the reduplicant as the base. If it is stipulated that the reduplicants in (5) are prefixes, rather than suffixes, then the correct base under the single-side definition will be chosen. But then there is a larger problem: there is no way to tell, *a priori*, what the base is. There are cases where the reduplicant is positioned as it is in (5) but copies from the preceding segments.[2] This occurs in Dakota, as shown in (6).

(6) Dakota Negation (Shaw 1976)[3]

(a)	(i)	haske-šni long-NEG	'not long'
	(ii)	haska-<u>ska</u>-šni long-PL-NEG	'not long (pl)'
(b)	(i)	apʰe-šni strike-NEG	'didn't strike'
	(ii)	apʰa-<u>pʰa</u>-šni strike-REPEATEDLY-NEG	'didn't strike repeatedly'

This would mean that looking at the same morpheme order, ROOT-RED-SUFFIX, the definition in (1) would need to consider the reduplicant in Nahuatl to be a prefix because it copies from the following string of segments but would need to consider the reduplicant in Dakota to be a suffix because it copies from the previous string of segments. Thus, following the single-side definition, the reduplicant in these cases could not be considered to be a prefix or a suffix until it is seen from which side it copies. However, the single-side definition is meant to pick out the base with regard to the prefixal/suffixal status of the reduplicant. If the base cannot be picked out until what the reduplicant copies has been identified, the definition is circular. Under the proposed definition of the base this problem does not arise because the base is not identified by the prefixal or suffixal status of the reduplicant. Rather, the base is always defined as the entire output with the exception of the reduplicant; that is, all the material which is not part of the reduplicant morpheme (the realization of RED). So the base-reduplicant relation in the proposed view is the reduplicant-nonreduplicant relation. Adopting the proposed definition solves the puzzle presented by the Nahuatl and Dakota data, as it defines the base in the same way for both, in fact, all, languages.

Third, it is not clear how the single-side definition would be implemented in optimality theory (OT) (Prince and Smolensky 1993). In the single-side definition you need to know the prefixal/suffixal status of a reduplicant in order to know what the base is. A base-reduplicant (BR) constraint then, needs to have access to the information of whether the reduplicant in an output candidate is a prefix or a suffix in order to know what segments the reduplicant is to be judged with respect to. OT assumes that the prefixal/suffixal status of an affix follows from the ranking of alignment constraints relativized to the morpheme (i.e. ALIGN

[2] I use the term "copy" descriptively.
[3] There is a difference in the last vowel of the root in the unreduplicated and reduplicated forms due to an ablaut rule that underapplies in reduplicated forms (Shaw 1976).

(AFFIX$_1$, PRWD)-LEFT[4] and ALIGN (AFFIX$_1$, PRWD)-RIGHT) rather than from information in the input. So a morpheme has no prefixal or suffixal status in the input but is placed in the output by the ranking of relativized alignment constraints. Since whether a reduplicative morpheme is a prefix or a suffix depends on a constraint ranking, it is not clear how a BR constraint would have access to this information. It seems that a BR constraint would have to know the relative ranking of the alignment constraint(s) relativized to the reduplicant in order to know what to evaluate as the base. There is no mechanism in OT that can do this, and building one in would defeat the OT tenet that constraints are evaluated independently of each other. Under the proposed definition of the base any BR constraint will always evaluate the reduplicant with respect to everything else in the output, and so it does not need to 'know' anything.

I suggest that stipulation of a particular base is unnecessary as well as problematic. By adopting the proposed definition, that the base is always everything in the output with the exception of the reduplicant, the identification of the base is both simplified and incurs none of the problems encountered by the single-side definition.

2.1. A word about anchoring constraints

Given my hypothesis that the base is everything in the output with the exception of the reduplicant, a few more words need to be said about the anchoring of the reduplicant. Taking the base to be everything other than the reduplicant means that, for example, the infixed reduplicant in Samoan [a-lo-loʎfa] ('love') (Marsack 1962) does not satisfy ANCHORBR-LEFT as the left edge of the base is [a...]. However, under the single-side definition of the base, [loʎfa] would be taken to be the base and the reduplicant would satisfy ANCHORBR-LEFT. So under the proposed definition, only the left and right edges of the output (excepting the reduplicant) can be referred to by the constraints ANCHORBR-LEFT and ANCHORBR-RIGHT.

In addition to being anchored to the base, I assume that the reduplicant may also be anchored to constituents within the base. There are languages in which the reduplicant is anchored to the root, and languages in which the reduplicant is anchored to the stressed syllable. Broselow and McCarthy 1983:53 note that "reduplicative morphemes may be prefixed not only to morphological but to phonological constituent", i.e. the syllable bearing the main stress. Thus, in addition to constraints which anchor the reduplicant to the prosodic word, I assume ANCHOR (RED, STRESSEDSYLL) and ANCHOR (RED, ROOT). The Somoan example above is a case where we would assume that ANCHOR (RED, STRESSEDSYLL) is highly ranked. Such anchoring constraints also make it possible to correctly limit what is copied by the reduplicant since the proposed

---

4 I assume the standard definitions of all constraints referred to. See, for example, McCarthy and Prince 1999.

definition defines the entire output, excepting the reduplicant but including all other affixes, as the base.

3. Theoretical consequences

Adopting the proposed definition of the base leads us to expect a pattern of reduplication which is not predicted under the single-side definition. It is predicted that the reduplicant may stand in correspondence with both preceding and following segments, as both are part of the base. The single-side definition of the base predicts this not to be possible because the base is limited to either the segments preceding or the segments following the reduplicant. I am aware of several cases where copying from both sides potentially occurs, and these are discussed in section 3.1. A second consequence of adopting the proposed definition of the base is that any BR constraint will always be evaluated with respect to the entire output, excepting the reduplicant. I look at a case where the single-side definition of the base makes a different prediction in section 3.2.

3.1. Copying from both sides

There are several languages that have been analyzed as cases where the reduplicant copies both from the preceding and from the following segments. This would not be possible under the single-side definition of the base, but is expected under the proposed definition. One such case is "out-of-control" reduplication in the Salish language Spokane.

(7) Spokane out-of-control reduplication (Bates and Carlson 1998)

(a) č'ehek'^{w}	'uncovered'	č'eh-hek'^{w}-k'^{w}	'it suddenly became uncovered'
(b) p'araq'	'turned back'	p'r-raq'-q'	'it got turned back by accident'
(c) šilič	'turn'	šl-lič-č	'it got turned by accident'

A process of unstressed vowel deletion in the base accounts for the vowels missing from the bases in the examples in (7) (Bates and Carlson 1998). Under this analysis, the reduplicant is infixed, and copies its initial material from the preceding segment and its final material from the following segment. The single-side definition of the base is not compatible with this analysis, but the proposed definition is.

The reduplication patterns of Hausa and Mangarayi have also been analyzed as cases of copying from both sides. The data are given in (8) and (9) where the (i) forms show the reduplicant as copying from both sides and the (ii) forms label the reduplicant as copying only the from the following segments.

(8) Hausa (Leben 1977:425, Newman 1972:318) (nominal roots are given in parentheses following Davis 1988:317)

(a) baaki (bak) 'mouth' baak(+RED)+unaa 'mouths'
- (i) baak-unk-unaa
- (ii) baa-kun-kunaa

(b) tuduu (tud) 'high ground' tud(+RED)+unaa 'high ground'
- (i) tud-und-unaa
- (ii) tu-dun-dunaa

Hausa is analyzed by Davis, among others, as a case of copying from both sides (i.e. as being the (i) forms) because that would place the reduplicant at a morpheme boundary, as the segments to the reduplicant's left make up the root, and the segments to its right are a plural suffix. Davis further notes that if the reduplicant in Hausa were as is represented in the (i) forms, it would have all the characteristics of an interfix, defined by Dressler 1985 as a semantically empty morpheme placed between a stem and a suffix. The reduplicative morpheme in (8) does not add any meaning to a word, as the unreduplicated form, for example, [baakunaa] for (8a), has the same meaning. Therefore this form of the reduplicant in Hausa can be considered an interfix if it is analyzed as copying from both sides. If the reduplicant is analyzed as in the (ii) forms instead, it does not fall at the morpheme boundary, and instead it is strangely infixed to the left of the final segment of the root. Hausa, then, seems to be case where the reduplicant is best analyzed as copying from both sides, showing us that both sides should be considered to be the base.

Reduplication in Mangarayi is also often analyzed as a case of copying from both sides (Merlan 1982, Davis 1988, Jones 1997).

(9) Mangarayi (Merlan 1982:216)

(a) jimgan 'knowledgeable person
- (i) jim-gim-gan 'knowledgeable people'
- (ii) j-img-imgan

(b) jalwayi 'muddy'
- (i) jal-wal-wayi 'very muddy'
- (ii) j-alw-alwayi

The analysis involving copying from both sides can be seen to have the virtue of obeying ALIGN (MORPH, SYLLABLE), which demands that the edge of every morpheme fall at a syllable boundary. The forms in (i) do this whereas the forms in (ii) do not. How an OT analysis can capture the difference between the two reduplicant options is shown in (10).

(10) Mangarayi

/RED+jimgan/	ALIGN (MORPH, SYLLABLE)	LINEARITYBR
a. ☞ jim-gim-gan		*
b. j-img-imgan	*!*	

Candidate (b) incurs two violations of ALIGN (MORPH, SYLLABLE) because both the left edge and right edge of the reduplicant are one segment off from a syllable boundary. Candidate (a) wins in this tableau because it satisfies ALIGN (MORPH, SYLLABLE), at the cost of violating LINEARITYBR. So although candidate (b) will win under the opposite ranking and there is no way to know for sure which option Mangarayi takes, it can be seen that it is possible to motivate and analyze a reduplicant which copies from both sides of its base. The single-side definition of the base does not predict the (i) forms in (9) to be possible. It is hard to see how the ranking in (10) could be ruled out. Therefore the single-side definition of the base faces the challenge of making sure a ranking never surfaces which would favor a reduplicant that copies from both sides.

3.2. MAXBR as base-minimizer

The second consequence of adopting the proposed definition of the base is that BR constraints will always evaluate the reduplicant with to respect everything else in the output. Reduplication in Nakanai has previously been analyzed in OT assuming the single-side definition of the base (Spaelti 1997) and I will show here how it works under the proposed definition of the base. Examples of reduplication in Nakanai are given in (12). (Unreduplicated meanings not available.)

(12) Nakanai reduplication (selected forms) (Johnston 1980)

(a)	vigilemuli	vigile-muli-múli	'story'
(b)	karusu	ka-rusu-rúsu	'ribs/battens'
(c)	lua	i–la-lúa	'two by two'
(d)	pita	pa-pîta	'muddy'

The stress is penultimate in Nakanai and so it can be deduced that the reduplicant is anchored to the left of the stressed syllable (ANCHOR (RED, STRESSSYLL)-LEFT is highly ranked). Spaelti insightfully unites these forms and others into one pattern of reduplication, noting that if the first vowel of what he takes to be the base (the segments following the reduplicant) is less sonorous than the second, that vowel is skipped and the next vowel is copied into the reduplicant instead. When the two vowels are equal in sonority, both are copied, as in the first two examples in (12). It is clear from the last two examples above that the reduplicant is the copy to the left and so we can deduce that the reduplicant in the first two forms is infixed rather than suffixed.

Spaelti, assuming the single-side definition of the base, takes the base to be only the segments that follow the reduplicant. He argues that the reason a reduplicant is infixed is to make the base smaller, and thus to better satisfy MAXBR. His tableau for the example in (12a) is reproduced in (13).[5]

(13) Spaelti 1997:162

/RED+vigilemuli/	ALL-FT-RT	MAXBR	ALIGN (RED, PRWD)-LT
a. ☞ vi(gile)(muli)(múli)	**/****		vigile
b. (vigi)(lemu)le(múli)	***/*****!	li	vigi
c. vi(gile)(gile)(múli)	**/****	muli!	vi
d. (vigi)vi(gile)(múli)	**/*****!	lemuli	

Spaelti is able to rule out candidate (c) in (13) because he has defined the base of reduplication to be the segments following the reduplicant, under the single-side definition of the base. Therefore the base is flexible, depending on the position of the reduplicant: candidate (c) incurs four violations of MAXBR whereas candidate (a) incurs none, even though both copy the same number of segments. But it is unclear why the constraint MAXBR is only sensitive to some of the segments in the output. In some way the segments following the reduplicant must be labeled as the base. Where and how did this labeling happen? The base as defined in the single-side definition is not labeled in the input, as the placement of the reduplicant as a prefix or a suffix is derived from the evaluation. If the base is labeled in the output, there must have been an independent operation that applied the single-side definition to each of the output candidates and thus marked some string of segments in each as the base. However, no operation has been claimed to act on the output candidates in OT except Evaluation. Therefore, the labeling of the base must come from the constraints themselves.[6] But then the problem this causes, as discussed in section 2, is again an issue: MaxBR in the above tableau must be able to see that ALIGN (RED, PRWD)-LEFT is higher ranked than ALIGN (RED, PRWD)-RIGHT (not shown in (13) but assumed to be lower-ranked). Knowledge of the ranking of other constraints is not a property OT constraints have. Therefore, the single-side definition of the base runs afoul of the limits of the OT framework. In the proposed definition, no stipulation of a particular base is needed, and there is nothing to implement in OT because base-reduplicant

[5] To help the winner stand out, I don't shade any cells of the winning candidate in my tableaux.

[6] There is another possibility, and that is that base is randomly specified in the output candidates. This alternative is examined by Haugen, Hicks Kennard and Kennedy 2002, and is argued against because it generates many more output candidates and does not add any simplification or elegance to the theory of reduplication.

correspondence constraints evaluate the reduplicant with respect to everything else in the output. The base is defined on the output to be the material that is not the reduplicant. Therefore, how BR constraints are evaluated does not change from one candidate to the next, as it does with MAXBR in (13). This seems a very desirable result of adopting the proposed definition.

Given the definition of the base I am proposing, all the candidates in (13) perform equally on MAXBR as all copy four out of ten segments of the base, thus each incurring six violations of MAXBR. And, given the rankings in (13), candidate (c) will incorrectly win if MAXBR is evaluated according to the proposed definition because it better satisfies the alignment constraint, as shown in (14).

(14) Spaelti's tableau with MAXBR evaluated under proposed definition of the base

/RED+vigilemuli/	ALL-FT-RT	MAXBR	ALIGN (RED, PRWD)-LT
a. vi(gile)(muli)(múli)	**/****	vigile	vig!ile
b. (vigi)(lemu)le(múli)	***/*****!	vigili	vigi
c. ☞WRONG! vi(gile)(gile)(múli)	**/****	vimuli	vi
d. (vigi)vi(gile)(múli)	**/*****!	lemuli	

It is ANCHOR (RED, STRESSEDSYLL)-LEFT which will make the choice of (a) over (c), as the left edge of the reduplicant in (c) is anchored four segments away from the left edge of the stressed syllable, as shown in (15). Thus, by appealing to anchoring constraints, we can rule out candidate (b) without appealing to a flexible base.

(15) ANCHOR (RED, STRESSEDSYLL)-LEFT rules out (b),(c) and (d) in (14)

/RED+vigilemuli/	ANCHOR (RED, STRESSSYLL)-LEFT	ALIGN (RED, PRWD)-LEFT
a. ☞ vi(gile)(muli)(múli)		vigile
b. (vigi)(lemu)le(múli)	l!e	vigi
c. vi(gile)(gile)(múli)	g!ile	vi
d. (vigi)vi(gile)(múli)	v!igile	

4. Conclusion

I have proposed that the base in reduplication is the entire output, excepting the reduplicant, contra the standardly assumed single-side definition. I have shown that the assumption that the base is everything except the reduplicant has multiple desirable consequences. It is simpler than the single-side definition, in which the base changes depending on what kind of affix the reduplicant is. It is not subject to the same limitations and problems that the single-side definition has. Finally, it makes an arguably correct prediction that the reduplicant may copy material from both the segments preceding it and those following it, something that is not predicted to occur under a definition where the base is taken to be one side or the other of the reduplicant. As MAXBR can no longer be used as a base-minimizer under the proposed definition of the base, I have offered another analysis for one case where this was employed in the literature. Defining the base as the entire output excepting the reduplicant has allowed me to account for the data examined without any stipulation of a particular base, and with a unified definition of the base which requires no special implementation in OT.

References

Bates, Dawn, and Barry F. Carlson. 1998. Spokane Npoqíniščn syllable structure *Salish Languages and Linguistics: theoretical and descriptive perspectives*, 99-123. Mouton de Gruyter.

Broselow, Ellen and John McCarthy. 1983. A theory of internal reduplication. *Linguistic Review* 3:25-88.

Chisholm, Matt. 2001. Nominal reduplication in Nahuatl. Ms.,University of California, Santa Cruz.

Davis, Stuart. 1988. On the nature of internal reduplication. In Michael Hammond and Michael Noonan (eds.) *Theoretical Morphology: Approaches in Modern Linguistics,* 305-323. San Diego: Academic Press, Inc.

Dressler, W. 1985. On the predictiveness of natural morphology. *Journal of Linguistics* 21:321-337.

Hale, Kenneth, and Abanel~Lacayo Blancho. 1989. Diccionario Elemental del Ulwa (Sumu Meridional). Center for Cognitive Science, MIT, Cambridge. *Non Vidi.*

Haugen, Jason, Catherine Hicks Kennard, and Robert Kennedy. 2002. The basis for bases: assigning reduplicative bases with alignment constraints. Handout from talk given at LSA 2002, San Francisco, CA.

Healey, Phyllis M. 1960. An Agta Grammar. Philippine Branch, Manila: Summer Institute of Linguistics.

Kager, René. 1999. *Optimality theory.* Cambridge: Cambridge University Press.

Johnston, Raymond L. 1980. Nakanai of New Britain: The Grammar of an Oceanic Language. *Pacific Linguistics*: Series B-70.

Jones, Caroline. 1997. Contiguity under infixation: Mangarayi reduplication. Handout from a talk given at UMass Amherst, May 11, 1997.

Leben, W. 1977. Doubling and reduplication in Hausa plurals. In A. Juilland (ed.) *Linguistic Studies Offered to Joseph Greenberg*, 419-431. Anma Libri.

Marsack, C.C. 1962. *Teach Yourself Samoan*. London: Hodder and Stoughton.

McCarthy, John and Alan Prince. 1993a. Prosodic Morphology I: constraint interaction and satisfaction. Technical Reports of Rutgers University Center for Cognitive Science, no. 3.

McCarthy, John and Alan Prince. 1993b. Generalized Alignment. In G.E. Booij and J. van Marle (eds.)*Yearbook of Morphology*, 79-153. Kluwer.

McCarthy, John and Alan Prince. 1999. Faithfulness and identity in prosodic morphology. In René Kager, Harry van der Hulst, and Wim Zonneveld (eds.) *The Prosody-Morphology Interface*, 218-309. Cambridge University Press.

Merlan, Francesca. 1982. *Mangarayi*. Lingua Descriptive Studies, Volume 4. Amterdam: North Holland Publishing Company.

Newman, P. 1972. Syllable weight as a phonological variable. *Studies in African Linguistics* 3:301-323.

Prince, Alan and Paul Smolensky. 1993. Optimality theory: constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder.

Shaw, Patricia A. 1976. Dakota phonology and morphology. PhD Dissertation, University of Toronto. Published 1980, New York: Garland Publishing.

Spaelti, Phillip. 1997. Dimensions of variation in multi-pattern reduplication. PhD Dissertation, University of California, Santa Cruz.

Sullivan, Thelma. 1988. *Compendium of Nahuatl Grammar*. Salt Lake City, Utah: University of Utah Press.

Lexical Representation in Korean: Evidence from the Tip of the Tongue Phenomenon[1]

SUNYOUNG HONG
University at Buffalo, The State University of New York

1. Introduction

The 'tip of the tongue (henceforth TOT) phenomenon' is a natural occurrence in speech production planning. The TOT phenomenon refers to the situation in which speakers cannot recall the complete phonological form of an intended word although they are sure that the word is known. In this state, speakers often retrieve a partial phonological form or other information related to the target word while they are trying to recall it. The information that is produced during word search functions as a window to lexical representation and language processing. Most TOT studies have been done in English or German, and there are no published TOT studies based on Korean. The current study attempts to answer questions regarding lexical representation in Korean, using the TOT experimental paradigm. Below are the specific research questions of this study:

(1) Research questions of the current paper
a. What phonological information is stored in lexical representations?
b. Do specific underlying phonological representations posited for Korean words have any cognitive validity?
c. Does the origin of words influence the structure of lexical representations in the mental lexicon? If yes, how does this affect the lexicon?
d. What is a proper model of lexical representation and retrieval for Korean?

In section 2, I give an overview of previous TOT studies with respect to current research questions. In section 3, I describe the TOT experiment designed for the current study. In section 4, I report the results in terms of aspects of the retrieved information. In section 5, I address the research questions listed in (1) above by elucidating the TOT data and relating them to models of lexical representation. Finally, in section 6, I provide a conclusion and make suggestions for future research.

---

[1] I would like to express my gratitude to Dr. Jeri J. Jaeger for her insightful advice.

2. Review of Tip of the Tongue studies

Among studies on the TOT phenomenon, the earliest experiment, considered the prototype for future experiments, was performed by Brown and McNeill (1966). In their experiment, the Ss were presented with 49 definitions of low-frequency words and then asked to recall the target words. At that time, those who could not think of the words but felt sure that they knew them were requested to write down certain aspects of the target words or to list other words that struck them as similar. Ss retrieved phonological details such as the number of syllables, the initial or final letter of a given word, the syllabic stress, etc. Brown and McNeill concluded that the more easily retrieved properties of low-frequency words carry more information than other properties and are thus paid more attention in speech perception. They suggested a model of lexical representation and access; their "faint-entry" theory assumes that a lexical item is represented in a multiply indexed format in the mental lexicon.

Koriat and Lieblich (1975) found that Ss often correctly retrieved final letters of target words when they simply guessed information about target words, which was interpreted as signaling a "correlation between semantic and structural information in word-final positions."

Jones and Langford (1987) argued that words being generated for production are first activated in the semantic component of the lexicon, with the activation then spreading to the phonological component, as mentioned in Morton's modified logogen model (Morton 1979, Jones, 1985). The Transmission Deficit Model of TOTs[2] proposed by Burke et al. (1991) also distinguished the semantic system from the phonological system. They posited several nodes within each system; from top to bottom: propositional node, lexical nodes, syllable nodes, phonological nodes, and feature nodes.

3. Methodology

3.1. Subjects

A total of 20 native Korean speakers (10 males and 10 females) who had been recruited from University at Buffalo participated in the experiment. Their ages ranged from 23 to 40. All Subjects (henceforth Ss) had a high degree of education; 14 spoke the Seoul dialect of Korean, and 6 spoke the North Kyungsang dialect.

3.2. Materials

The materials consisted of definitions of 60 Korean words. They include 15 native Korean (henceforth NK) words, 25 Sino-Korean (henceforth SK) words[3] and 20 recent loan (henceforth RL) words. To increase the probability of the TOT state,

[2] This model was developed from the Node Structure Theory suggested by MacKay (1987), which is an interactive activation model of language production and perception.

[3] SK words refer to words borrowed from Chinese several hundred years ago. Since they have been used for a very long time, they are not considered RL words.

the materials consisted of mid-frequency words[4]. They were all uninflected nouns including both monomorphemic and multimorphemic nouns. The number of syllables of the words ranged from 2 to 5.

3.3. Procedures

The Ss were tested individually and the whole session was recorded. Each subject was presented with 60 definitions in a random order. As soon as they heard each definition, the Ss were directed to say the target word. If they felt that they did not know a word, they said so and skipped over the word. However, if they felt that they had gone into the TOT state, they were asked to both speak out loud and write down everything that occurred to them about the word; they were allowed at most three minutes to find the target word. After three minutes, I read out the target word, and then the Ss in the TOT state were asked to report whether it was the word they were searching for. If they reported it was not the word, they were asked whether they had been searching for a different word, or had simply guessed several aspects of the target word.

4. Results

4.1. General results

The responses of each S to each definition were classified into the six categories given in (2) below, which are based on the five categories defined by Kohn et al. (1987):

(2) Types of Ss' search state

a. Don't know: S felt that s/he did not know the target word. This was verified when the word was read out by the experimenter.

b. Guess: S did not search for any particular word, but provided some aspects that they guessed.

c. Negative TOT: S was searching for an incorrect target word.

d. Positive TOT: S was searching for the correct target word, but did not produce the target word.

e. Final Success: S finally produced the target word within the allotted period of time after experiencing the TOT state.

f. Immediate Success: Ss recalled the target word immediately after hearing the definition.

I counted how many Ss declared themselves to be in one of the four 'search states': guess, negative TOT, positive TOT, final success. The table in (3) below summarizes Ss' responses in terms of search state.

---

[4] Because there are currently no documented references concerning frequency of Korean words, I had two Korean linguists screen the original word list consisting of 129 words and then judge word-frequency based on their language usage and linguistic knowledge. I only included words that they judged as mid-frequency words.

(3) Distribution of search states of 20 subjects for 60 target words

	Guess	Negative TOT	Positive TOT	Final success	Total
No. of cases	15	55	234	92	396

In this study, I only included the responses that subjects produced when they were in positive TOT state or final success state to ensure that subjects were looking for correct target words. The raw data collected from my TOT experiments are categorized according to the common attributes of Ss' responses. The categories of responses are defined in (4) below.

(4) Categories of responses according to their attributes

a. Number of syllables: Ss attempted to retrieve how many syllables the target words had.

e.g. When the target word /kol.mu/ 'thimble' induced the TOT state, seven Ss recalled that it was a disyllabic word.

b. Words of similar sound (henceforth SS word): Ss produced complete words during the search, which did not completely match the target word but were phonologically similar to the target words[5]. When they produced words, they said that the words were either right or fairly similar to target words in sounds. SS words could be either a real word or a non-existent word.

e.g. When the word /pok+c^{h}ɛ/[6] 'fee for having one's fortune told' induced the TOT state, three Ss recalled a real word /pok.cu/, which is phonologically related to the target word in that it has the same number of syllables and the first syllable is identical.

c. Phonological fragments: Ss attempted to retrieve part of a word: a single consonant (henceforth C), a single vowel (henceforth V), sub-syllabic unit, and syllable.

e.g. When the word /sin+mun+ko/ 'drum which people hit to complain of an injustice' induced the TOT state, two Ss recalled that it ended in /ko/.

d. Other phonological information: Ss attempted to retrieve phonological information other than specific C, V, sub-syllabic fragment, and syllable.

e.g. When the word /sa+li/ 'relic of the Buddha' induced the TOT state, a S said that there was no aspirated sound in the target word.

e. Non-SS word: Ss produced words, but these words were not phonologically similar to the target words. In some cases, Ss immediately mentioned that the words were incorrect and phonologically unrelated to the targets.

e.g. When the word /tɛ.nim/ 'traditional pant-leg ties' induced the TOT state, a S recalled a word /ko.ɾɨm/ 'breast-tie', which is semantically related in that both are the names of ties attached to Korean traditional dress.

[5] The criterion for SS word as opposed to non-SS word was that the retrieved words should match at least one syllable or two separate segments in target words.

[6] In this paper, I transcribe a palatal affricate as /c/ for typographic simplicity.

f. Other non-phonological information: Ss attempted to retrieve non-phonological information such as word origin class regarding target words.

e.g. When the word /ə+joŋ/ 'government use' induced the TOT state, a S mentioned that the target word was a Sino-Korean word.

Reflecting the specific categorizations defined in (4) above, the table in (5) shows how many responses there were in each category regardless of the correctness of the responses. Note that in many cases, a single S provided more than one responses when s/he experienced the TOT state.

(5) Distribution of responses according to categories

	No. of responses	% of responses
a. Number of syllables	232	34%
b. SS word	155	23%
c. Phonological fragments	148	21%
d. Other phonological information	22	3%
e. Non-SS word	101	15%
f. Other non-phonological information	27	4%
Total	685	100%

As shown in (5), the distribution of the response categories is as follows: number of syllables>SS words>phonological fragments>non-SS word>others.

4.2. Number of syllables

Responses involving the number of syllables are summarized in (6) below.

(6) Total responses involving the number of syllables

Actual numbers	Responses 1	2	3	4	5	Total	Correct	Mode	Mean	Difference[7]
2	1	**80**[8]	18	1	0	100	80%	2	2.19	+0.19
3	0	6	**60**	8	2	76	79%	3	3.08	+0.08
4	0	2	15	**29**	6	52	56%	4	3.75	-0.25
5	0	0	2	1	**1**	4	25%	3	3.75	-1.25
Total	1	88	95	39	9	232	73%			

The table in (6) above shows that as the number of syllables in target words increases, accuracy of responses drops. In order to compare recall of the number of syllables for NK, SK and RL words, I separated the figures into three groups according to word origin. In NK words, the number of syllables was correctly retrieved 63% of the time; in SK words, 85% of the time; in RL words, 63% of

[7] 'Difference' means difference between means and actual number of syllables.
[8] The numbers in boldface in the table are the number of correct responses.

the time. To see if there is an association between word origin class and tendency of correctness in responses, a two-way chi square test was performed. The statistical result tells us that there is an association between them ($\chi^2(2)=14.897$, $p<.001$). Specifically, the number of syllables is more accurately retrieved in SK words than in NK or RL words.

4.3. Phonological units involved in fragment

I examined what kinds of units were involved in fragments that subjects retrieved during the TOT state. The table in (7) presents the results of the analysis.

(7) Types of phonological units regardless of location

		Single C	Single V	Body	Syllable	Total
Total	Attempt	10 (6%)	6 (3%)	7 (4%)	159 (87%)	182 (100%)
	Correct	10 (7%)	6 (4%)	7 (5%)	120 (84%)	143 (100%)

As seen in (7) above, the phonological units retrieved were Syllable>Single C>Body (CV-)>Single V. Interestingly, the unit of rhyme (-VC) was not involved.

4.4. SS word

I have investigated the extent to which the number of syllables in SS words matched the number of syllables in target words in comparison with the case of non-SS words. The table in (8) presents the result.

(8) Match in the number of syllables in SS words and non-SS words

	Same number	Different number	Total
SS words	117 (75%)	38 (25%)	155 (100%)
Non-SS words	56 (56%)	44 (44%)	100 (100%)

The differences between same number of syllables and different number of syllables are significantly associated with the category of SS words and non-SS words ($\chi^2(1)=10.577, p<.001$).

The table in (9) shows the number of complete matches of whole syllables in the same location in target words and SS words.

(9) Syllable match in SS words

	Initial syllable	Medial syllable	Final syllable	Total
NK	8 (25%)	5 (16%)	19 (59%)	32 (100%)
SK	18 (21%)	15 (17%)	53 (62%)	86 (100%)
RL	25 (33%)	36 (48%)	14 (19%)	75 (100%)
Total	51 (26%)	56 (29%)	86 (45%)	193 (100%)

The figures in (9) above can only be interpreted with reference to the possible number of matches. In order to calculate this figure, I counted the number of initial, medial, and final syllables in target words. The count of possible matches is given in (10) below, by word origin.

(10) Possible numbers of syllable matches, by syllable location and by word origin

	Initial syllable	Medial syllable	Final syllable	Total
NK	15 (39.5%)	8 (21%)	15 (39.5%)	38 (100%)
SK	25 (37.5%)	17 (25%)	25 (37.5%)	67 (100%)
RL	20 (28.5%)	30 (43%)	20 (28.5%)	70 (100%)
Total	60 (34.5%)	55 (31%)	60 (34.5%)	175 (100%)

Looking first at NK words, it can be seen that the final syllables were recalled more often than would be expected by frequency of occurrence (59% match vs. 29.5% occurrence). Similar figures were found for SK words, with an even more pronounced effect of final-syllable recollection (62% match vs. 37.5% occurrence). A different pattern was found for the RL words. Initial syllables rather than final syllables matched more often than would be predicated by the frequency figures (initial=33% match vs. 28.5% occurrence).

4.5. Phonological fragments

I have examined the responses where Ss in the TOT state produced parts of words. The table in (11) below summarizes the cases in which an entire syllable in the response matched a syllable in the target word.

(11) Syllable match regardless of correctness of locations in SK and RL words[9]

	Initial	Medial	Final	Total
SK	12 (29%)	6 (15%)	23 (56%)	41 (100%)
RL	12 (63%)	7 (37%)	0 (0%)	19 (100%)

With SK words, final syllables were much more frequently retrieved correctly. Conversely, initial syllable match was most predominant in RL words. This finding in SK and RL words corresponds to the result displayed in (9) in that the final syllable is more correctly retrieved in SK words and the initial syllable is more correctly retrieved in RL words.

4.6. Other

There were other types of responses involving other phonological information illustrated in (4d), non-SS words in (4e), and other non-phonological information in (4f). However, they did not fall into any particular patterns, so will not be further discussed in this paper.

[9] Note that as there were only 5 cases of syllable match in total in NK words, NK words are not considered here.

5. Discussion

5.1. Lexical representation in Korean

This section addresses the first research question in (1) by examining Ss' responses regarding phonological information. As presented in section 4, responses that Ss provided while in the TOT state included the number of syllables, SS words, phonological fragments, phonological information such as categorical or positional information of segments, non-SS words, and other non-phonological information. Based on these results, I will discuss how robustly these various linguistic properties are represented in the phonological form of the mental lexicon.

5.1.1. Syllables

This section concerns recollection of the number of syllables. From prior studies, it is known that, in English, the number of syllables is one piece of phonological information that is usually easily retrieved in the TOT state. The current study shows that this is also true for Korean. As presented in (5), one third of the total responses were concerned with the number of syllables in the target word. Not only did Ss frequently try to recall the number of syllables, but they also recalled this figure quite accurately. Furthermore, the argument that the number of syllables is robustly represented in the lexical entries is also supported by the fact that SS words (75%) had the same number of syllables as target words more often than non-SS words (56%).

Strictly speaking, however, the number of syllables was correctly retrieved 85% of the time in SK words, while it was correctly retrieved 63% of the time in NK and RL words, respectively. This result suggests that syllables of SK words are more robustly represented in the mental lexicon than those of NK and RL words. In other words, it appears that SK words are clearly represented syllable by syllable in the mental lexicon, allowing speakers to easily count the number of syllables in their minds even when they cannot recall the complete phonological form. The most likely explanation for this involves the morphological structure of SK words; each syllable exactly corresponds to an individual morpheme in SK words. For example, the disyllabic word /p^{h}uŋ+kjəŋ/ 'wind-bell' consists of two morphemes; the first syllable corresponds to a morpheme /p^{h}uŋ/ 'wind' and the second syllable corresponds to a morpheme /kjəŋ/ 'hand-bell.' Based on the fact that each syllable matches an individual morpheme in SK words, it can alternatively be said that SK words are represented morpheme by morpheme in the lexicon and speakers count the number of morphemes in their minds rather than the number of syllables. However, since Ss quite accurately recalled the number of syllables of NK and RL words as well as SK words, I would argue that syllables of words are represented in the lexicon regardless of word origin classes. In addition to that, I argue that the unique word structure of SK words contributes to frequent and correct retrieval of the number of syllables.

Another piece of evidence for the representation of syllables comes from analyzing phonological units involving fragments. The most frequent unit involving fragments that were correctly retrieved was syllable (84%). Hence, one could argue that Ss were more likely to access a node for an entire syllable than individual segments, or body, or rhyme.

Finally, evidence for syllables is found in the deletion of syllables in retrieved words. Ss often retrieved partial phonological forms while leaving out some segments. When phonological units were deleted, these units tended to be a syllable in the word's underlying form. For instance, a S produced [k^he.sɨ.c^hɨ] in search of word /k^he.sɨ.t^hə.ne.c^hɨ/ 'castanets', but dropped the third and fourth syllables.

I have shown that the syllable is a predominant unit in phonological representations in the Korean mental lexicon and so phonemes are grouped by syllables at some level in the phonological representations. In addition, words are hierarchically organized according to syllable level and individual segment level. This conforms to one of the arguments proposed by Fay and Cutler (1977) to some extent. Fay and Cutler argued from malapropism data that "the major partitioning of the dictionary seems to be by number of syllables.... (p. 516)"

5.1.2. Syllable structure

In order to investigate syllable structure consistency, SS words of the same number of syllables were analyzed. Target words and SS words in underlying forms if available were compared by syllables from left to right. The calculation system was as follows: common syllabic slots that ware occupied in both target and SS word were counted and then divided by the number of all possible syllabic slots. The syllable structure consistency between SS words and target words were 83.4%, while the consistency between non-SS words and target words were 69.9%. This difference was significant ($\chi^2(1)$=4.7, p<.03). These figures argue that at least the syllable structure slots of 'onset', 'glide,' 'nucleus' and 'coda' are represented in phonological representations.

5.1.3. Segments

Ss frequently attempted to recall individual segments of target words and also often produced SS words where some of segments are identical with those of targets. There is no doubt that individual segments are stored in phonological representations.

5.1.4. Prosody

There were no responses involving prosodic information of targets in Korean, as opposed to English where people often recalled the location of primary stress of target words or produced words that had the primary stress on the same syllable as the target word. This suggests that in Korean prosodic information is either not represented in the lexicon or at least not strongly represented in the lexicon.

5.2. Cognitive validity of underlying form posited for Korean words

Concerning the second research question about the cognitive validity of the underlying forms, one could argue that the underlying forms are too abstract and that speakers actually store forms more similar to the surface forms. To resolve this controversy, I have analyzed the SS words and the phonological fragments only when there was a difference between the posited underlying forms of the target words and the surface forms of the responses to examine cognitive validity of underlying forms. Compare the targets with the responses in (12-14).

(12) Phonological process involving the target: Tensification

a. target	/nak+**kwan**/	[nak.**k'wan**][10]	'painter's sign and seal'
b. fragment	**[kwan]**		

(13) Phonological process involving the target: Place assimilation

a. target	/**in**+kam/	[**iŋ**.gam]	'registered seal impression'
b. SS word	[kwa.**nin**] ~ /kwan.in/[11]		'official seal'

(14) Phonological process involving the target: Palatalization

a. target	/**sin**+mun+ko/	[ʃ̆im.muŋ.go]	'drum which people hit to complain of an injustice'
b. SS word	[**səŋ**.muŋ.go]		nonce word

Responses in (12-14) above showed that underlying forms posited for target words correspond to the Ss' phonological representations, i.e., the underlying forms have cognitive validity.

5.3. Effects of word origin on syllable retrieval

In this section, I discuss how the word origin affects syllable retrieval, which is the third research question. As seen in (11), the final syllable was most correctly retrieved in SK words, while the initial syllable was more correctly retrieved in RL words. For example, when the target word /c^{h}ən+toŋ+səl/ 'Geocentricism (universe+movement+theory)' induced the TOT state, Ss often recalled the final syllable /səl/ 'theory' only. This is best explained by the fact that final syllables in most SK multimorphemic nouns represent the superordinate component of the words. Therefore, it is argued that a syllable representing superordinate component of a word may be robustly represented in the mental lexicon, which suggests that there is more than one way to arrange phonological specification in the phonological representation as opposed to the left-to-right listing hypothesis argued by Fay and Cutler (1977)[12].

---

[10] The apostrophe as in [nak.k'wan] traditionally represents a tense sound in Korean.

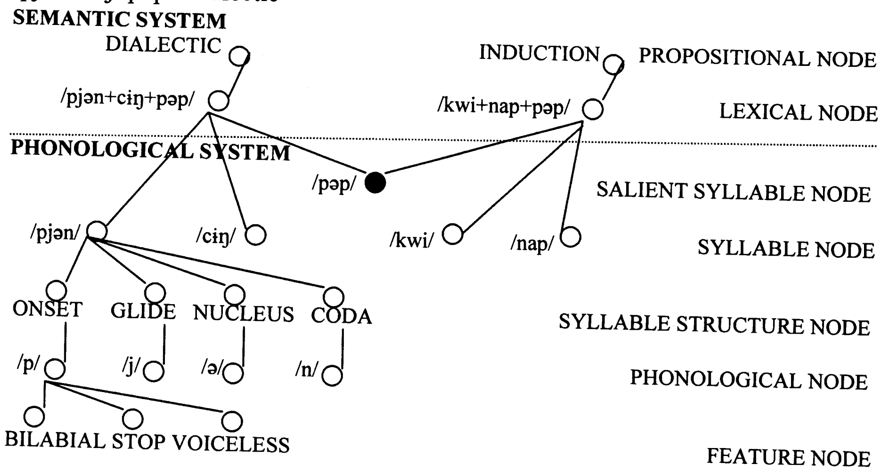
[11] The syllabification of the SS word /kwan.in/ is based on how Ss wrote down the word in Korean.

[12] Fay and Culter (1977) hypothesized that in the mental dictionary, "words are arranged by phonemic structure, in a left-to-right manner" and "the words that begin with the same phoneme are listed together, those that have the same second phoneme form a subcategory of that class, and so on."

5.4. Model of lexical representation and retrieval in Korean

Based on the findings of the TOT experiment in Korean, I propose a model of lexical representation and retrieval for Korean which was developed from the Transmission Deficit Model proposed by Burke et al. (1991) and Rastle and Burke (1996). As an example, the figure in (15) below illustrates the lexical representation of a SK word.

(15) Hypothetical lexical representation of the multi-morphemic SK word /pjən+cɨŋ+pəp/ 'dialectic'[13]



It should be noted that the final syllable is connected to a syllable node, illustrated with black color, which I call the 'salient syllable node.' This syllable is readily retrieved because of the meaning structure discussed in section 5.3. Furthermore, unlike Rastle and Burke's model for English, there is no information about stress available under syllable nodes. Also, Rastle and Burke grouped VC together forming a 'vowel group' at the level of phonological nodes, whereas I do not since there is no obvious reason to believe that the rhyme has special status in the Korean mental lexicon.

6. Conclusion

The current study attempted to resolve several issues regarding lexical representation in Korean by using the TOT method. This experiment confirmed that some phonological aspects such as the number of syllables are easily retrieved. However, it was argued that prosodic information is not strongly represented in the Korean lexicon. The data proved that underlying forms posited for target words have cognitive validity. I have also shown that phonological representations may partly differ according to word origin due to different

[13] /pjən/ stands for 'identification'; /cɨŋ/ for 'evidence'; /pəp/ for logic.

meaning structure. Finally, I proposed a model of lexical representation and retrieval for Korean on the basis of the findings.

For future research, it would be interesting to collect naturalistic TOT data and examine any significant difference between experimental data and naturalistic data. It would also be useful to test the TOT phenomenon cross-linguistically in order to determine the universality and specificity of lexical representation.

References

Brown, R. and D. McNeill. 1966. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5, 4: 325-337.

Burke, D. M., D. G. MacKay, J. S. Worthley, & E. Wade. 1991. On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language* 30: 542-579.

Fay, David, & Anne Culter. 1977. Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8: 505-520.

Jones, G. V. 1985. Deep dyslexia, imageability, and ease of predication. *Brain and Language* 24: 1-19.

Jones, G. V. and S. Langford. 1987. Phonological blocking in the tip of the tongue state. *Cognition* 26, 2: 115-122.

Kohn, S. E., A. Wingfield, L. Menn, H. Goodglass, J. B. Gleason, and M. Hyde. 1987. Lexical retrieval: the tip of the tongue. *Applied Psycholinguistics* 8, 3: 245-266.

Koriat, A., & I. Lieblich. 1975. Examination of the letter serial position effect in the "TOT" and the "don't know" state. *Bulletin of the Psychonomic Society* 6: 539-541.

MacKay, D. G. 1987. *The Organization of Perception and Action: A Theory for Language and Other Cognitive Skills*. New York: Springer-Veerlag.

Morton, J. 1979. Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.) *Processing of Visible Language* 1: 259-268.

Rastle, K. G. and D. M. Burke. 1996. Priming the tip of the tongue: effects of prior processing on word retrieval in young and older adults. *Journal of Memory and Language* 35, 4: 586-605.

Sunyoung Hong
Department of Linguistics
University at Buffalo, The State University of New York
609 Baldy Hall
Buffalo, NY 14260

sh16@acsu.buffalo.edu

Markedness and the Phonetic Implementation of Tone in North Kyungsang Korean

WILLIAM J. IDSARDI & WOOHYEOK CHANG
University of Delaware

0. Introduction

North Kyungsang Korean (NKK) has a pitch-accent system similar to that of various Japanese dialects. A distinctively high pitch can occur in four different positions within the word: final, penultimate, initial,[1] or on the first two syllables (which we will call double). Controversy in analyses has arisen as to which of the positions is the unmarked one. N.-J. Kim (1997) proposes an Optimality Theory (OT) analysis in which the default location of high tone is on the penultimate syllable of a word; other locations are lexically marked with pre-linked high tones. In contrast, S.-H. Kim (1999b) proposes a Simplified Bracketed Grid (SBG) metrical analysis in which the final position of a word is the default position; other locations are lexically marked with foot boundaries. The goal of this study is to employ phonetic evidence to resolve this debate.

We start with a basic observation that marked phonological elements tend to have more extreme values along phonetically measurable scales. A simple relevant case is the Korean stop system and the phonetic scales for closure duration (CD) and voice onset time (VOT). Although CD and VOT vary depending on the position in the word (between for example word-initial and intervocalic positions) the relative ranks stay constant. Thus, aspirated stops have longer VOTs, even though the plain stops are moderately aspirated word-initially. Likewise, the CD of tense and aspirated stops is longer than that of plain stops, even though the CD is reduced in word-initial position. Without committing ourselves at this point to an interpretation of such phenomena, we will use this observation to construct a test for the relative markedness of the tones. That is, we should expect to find that *marked* tones are phonetically *more* prominent than unmarked ones. In addition, the two contrasting theories make opposing predictions about the relative pitch values for tones in different positions. Under

[1] The Daegu dialect discussed here seems to be undergoing a change in which the initial tones are being replaced by double tones.

the OT analysis, the stem-final tone is marked in the input. Thus, the F_0 of the final tone in stems is predicted to be higher than the one of the non-final tone in stems (final H $\geq$ non-final H). In the SBG analysis, the stem-final tone is unmarked, whereas the non-final tone in stems is marked. Crucial to this analysis, the difference between the two high tones lies in the type of feet (open feet *vs.* closed feet). Since open feet are universally less prominent than closed ones (Idsardi 1994), final tone, which is in an open foot, should be lower than the non-final tone which is in a closed foot (final H < non-final H).

In order to test which prediction is correct, we conducted two phonetic experiments measuring the peak F_0 of matched sets of words. The first experiment shows that non-final accents have higher pitch than final accents. The second experiment examines the possibility that the difference observed in the first experiment is due to an effect of final lowering. The peak pitch of doubled high tones in two different environments was measured: word-finally (i.e., in disyllabic words) and non-word-finally (i.e., in trisyllabic words). The results show that there is no significant difference between the two means, which indicates that there is no general process of final lowering operating in NKK.

We therefore conclude that NKK penultimate tones are phonetically higher than final tones, and thus that the penultimate position is the marked one, consistent with the SBG analysis.

1. The tonal patterns of NKK tone

NKK has a pitch accent system in which each phonological word has a single high tone. There are four contrasting locations of a high tone (i.e., final tone, non-final tone, initial and double tone) in lexical items, as illustrated in (1).

(1)	a.	Final tone:		[...H] pattern		
		satalí	'ladder'		namwú	'tree'
		palám	'wind'		kaúl	'autumn'
	b.	Non-final tone:		[...HL] pattern		
		hánul	'sky'		apéci	'father'
		pwulkasáli	'star fish'		yángpok	'suit'
	c.	Initial tone:		[H...] pattern[2]		
		myénuli	'daughter-in-law'		ácime	'aunt'
	d.	Double tone:		[HH...] pattern		
		kúlím	'picture'		mwúcíkay	'rainbow'
		sáép	'business'		síktáng	'restaurant'

One piece of evidence for the relative markedness of the tones is the tonal changes observed in encliticized words. Stems of types (1b-d) maintain the

[2] As noted above, in the Daegu dialect the initial pattern (1c) is changing to the double tone type (1d). For this reason, initial tone patterns were excluded from the study.

position of the high tone of the stem in isolation, as shown in (2) and (3). Stems with final tone (1a), however, show a shift with consonant-initial enclitics, (4a).

(2) Non-final accented stem + enclitics
a. apéci 'father' + cocha 'even' (consonant-initial enclitic) → apéci-cocha
b. apéci 'father' + eykey 'to' (vowel-initial enclitic) → apéci-eykey

(3) Double accented stem + enclitics
a. kúlím 'picture' + chelem 'like' (consonant-initial enclitic) → kúlím-chelem
b. kúlím 'picture' + ulo 'with' (vowel-initial enclitic) → kúlím-ulo

(4) Final accented stem + enclitics
a. satalí 'ladder' + cocha 'even' (consonant-initial enclitic) → satali-cócha
b. satalí 'ladder' + eyse 'at' (vowel-initial enclitic) → satalí -eyse

This pattern is analyzed by S.-H. Kim as indicating that consonant-initial enclitics have an accent, that vowel-initial enclitics are pre-accenting, and that the surface tone appears on the first accented position of the word, as long as the stems with final tone are analyzed as unaccented, as discussed in the next section.

2. Two recent phonological analyses on the NKK tone system

The recent accounts of the NKK tone fall into two groups. Some previous analyses claim that the final accent is lexically marked and the penultimate accent is unmarked (G.-R. Kim 1988, N.-J. Kim 1993, 1997, and Kenstowicz and Sohn, 1997). Others (Y.-H. Chung 1991 and S.-H. Kim 1999a, b) propose that the final tone is the lexically unmarked one. S.-H. Kim's metrical analysis is summarized in (5). Lexical accents are represented with foot boundaries, and high tone is inserted on the final syllable of the first foot of the word. A difference then arises between *closed* feet (5b) – those with both foot boundaries – and *open* feet (5a). Idsardi (1994) argues that closed feet are stronger, and so in the present context we predict that the high tone in (5b) will be higher than that in (5a).

(5)

	a. Final accented stem	b. Penult accented stem
Underlying Rep.	x x x	x x) x
Surface Rep.	H | (x x x	H | (x x) x

In N.-J. Kim's (1997) OT analysis, the penult accented stems as analyzed as toneless (unaccented), while final accents are lexically represented with high tone. As S.-H. Kim (1999a) points out, the details of the analysis are not consistent with the principle of Lexicon Optimization in OT, which would select an alternative analysis in which all tones are lexically represented, minimizing the discrepancies between input and output representations. In such an analysis none of the stems would be unaccented. Therefore, N.-J. Kim's analysis predicts either that final accents should have higher pitch than non-final ones or (with Lexicon Optimization) that they should be equivalent.

Although the two accounts differ radically in the mechanisms and computations employed, we can reduce the controversy to a question of lexically marked versus unmarked tones. The two accounts differ on the relative status of final and penult tones and the difference of the accounts is attributed to the markedness in the underlying representation. As summarized in (6), the phonologically marked stems with regard to a high tone are different from each other depending on which phonological account we follow.

(6)

Analysis	Final accented stems	Penult accented stems
N.-J. Kim (OT)	Marked	Unmarked/Marked
S.-H. Kim (SBG)	Unmarked	Marked

3. Phonetic realization of NKK tone

So far, we have briefly reviewed two competing accounts of tone in NKK stems. Let us now attempt to establish which phonological account better correlates with phonetic evidence on the NKK tones. Based on the above observed phonological controversy, it is important to make some possible predictions for the following experiment, as illustrated in (7).

(7) a. Pitch predictions from N.-J. Kim's OT analysis

```
        final H  =  final H with vowel-initial enclitics  ≥  penult H
          H                    H                              H
          |         =          |                  ≥           |
      x  x  x              x  x  x  #  x  x               x  x  x
```

b. Pitch predictions from S.-H. Kim's SBG analysis

```
        final H        <  final H with vowel-initial enclitics, penult H
          H                    H                              H
          |         <          |                  =           |
      (x  x  x             (x  x  x #) x  x              (x  x)  x
```

Under reasonable assumptions about phonology phonetics interface (in particular a principle of *no markedness flip*) the two hypotheses make phonetically different predictions. From N.-J. Kim's OT analysis, we can infer two feasible predictions, as shown in (7a). First, all of the high tones, in terms of F_0, should be the same in the final accented stems, the final accented stem plus vowel-initial enclitic forms, and the penult accented stems. Second, the stem-penultimate high tone should be lower than the stem-final high tone in the isolated stems as well as in the encliticized words, since the final H marked from the beginning should be more prominent than the penult H specified only in the output (parallel to the moderate VOT observed in initial plain stops).

Following the SBG Theory-based analysis, we can, on the other hand, predict that the high tone linked to open feet (final accented stems) should be less prominent (lower F_0) than the high tone linked to closed feet (both penult accented stems and final accented stem plus vowel-initial enclitic forms), as summarized in (7b). Notice also that the SBG account produces representations which preserve markedness information. The right metrical bracket is preserved in the output, indicating the lexically marked positions. Thus, one way of interpreting higher pitch on marked tones is that the high tone itself raises the pitch target, and the right metrical boundary adds its own increment on top of that.

3.1. Experiment 1: comparison between final and penult tones

The purpose of this experiment is to test the predictions outlined in (7). If the subsequent experiment reveals that the F_0 of the stem-final high tones in the isolated stems and the encliticized words are as high as or even higher than the F_0 of the penult high tones, as illustrated in (7a), we would support N.-J. Kim's OT analysis. Conversely, if the result confirms the lower pitch value in the high tone of the final accented stems than in the high tone of the encliticized words and penult accented stems, as in (7b), we can support S.-H. Kim's SBG analysis.

3.1.1. Methods

Ten male native speakers of NKK participated in this experiment. Details about the participants are given in Chang (2002). The thirty words listed in (8) were selected as stimuli (e.g., ten words with a final high tone + ten words in their combination with vowel-initial enclitics + ten words with a penult high tone).

(8) Stimuli used in experiment 1

Final accent		Final accent + enclitic		Penult accent	
naympi	'pan'	naympi-eyse	'on a pan'	pinu	'soap'
satali	'ladder'	satali-eyse	'on a ladder'	talimi	'iron'
matang	'yard'	matang-eyse	'in the yard'	tangmyen	'noodle'
koa	'orphan'	koa-eykey	'to an orphan'	congali	'calf'
namwu	'tree'	namwu-eyse	'on a tree'	punmwuki	'spray'
mwuwu	'radish'	mwuwu-eyse	'in a radish'	twungwuli	'nest'
tongmwun	'fellow'	tongmwun-eykey	'to a fellow'	mwuncip	'anthology'
maum	'heart'	maum-ulo	'in the heart'	maumssi	'temper'
namwul	'greens'	namwul-eyse	'in greens'	mwultong	'bucket
seyang	'the West'	seyang-eyse	'in the West'	yangpok	'suit'

The syllables with high tone, underlined in (8), are the target syllables for the measurement of F_0. Unfortunately, F_0 can be affected by various other factors, such as the type of onset and the vowel quality. In order to at least partially control for these factors, the stimuli are composed of ten with the same target syllable. We did not include consonant-initial enclitics because the number of consonant-initial enclitics is limited, making it difficult or even impossible to construct triplets using the same target syllable as in the other items.

3.1.2. Procedure
The test words in (8) were written on index cards in Korean orthography. Each speaker read them, in random order, at a natural, comfortable speed. Two repetitions of the entire stack of cards were produced, resulting in a total 60 utterances (30 stimuli x 2 repetitions). The utterances were recorded onto a digital mini disc, using a high-quality microphone (Sony ECM-MS907) and mini-disc recorder (Sharp MD-MT821-A). The recorded words were then converted to WAV files on a computer at a 22 kHz sampling rate and 16 bit quantization. Next, F_0 contours were produced and measured by using Speech Analyzer 1.5, and the peak F_0 values of each word were collected.

3.1.3. Results
For each type of the stimuli 200 measurements were taken (10 subjects x 10 items x 2 repetitions). The mean F_0 and the standard deviation for each type are shown in (9). The final accent (9a) had the lowest mean pitch value.

(9)

Type	Number	Mean F_0	Standard deviation
a. Final accent	200	123.4 Hz	14.5 Hz
b. Final accent + enclitic	200	137.3 Hz	15.9 Hz
c. Penult accent	200	138.9 Hz	16.3 Hz

A repeated measures analysis of variance (ANOVA) shows that the difference in the mean F_0 values among the items is significant: $F(2,297) = 29.8$; $p < 0.0001$. A pairwise post-hoc Scheffé test, (10), shows that (9a) is significantly lower than (9b) and (9c), but that there is no significant difference between (9b) and (9c).

(10) Summary statistics for the Scheffé post-hoc test on F_0 for the three items

Contrasts	Mean Difference	P-Value
(9a) final *vs.* (9b) final + enclitic	13.9 Hz	< 0.0001
(9a) final *vs.* (9c) penult	15.5 Hz	< 0.0001
(9b) final + enclitic *vs.* (9c) penult	1.6 Hz	0.774

Since we do not know from the statistics whether this categorization reflects a uniform pattern across all subjects, we should examine the individual subjects' performances. For every subject, the pitch of the final accent in stems is consistently less than both of the other accent types, as shown in (11). On the contrary, no consistent patterns are found when the final accent in encliticized words is compared with the penult accent in stems. Only six out of ten subjects show a higher F_0 in penult position.

Due to the consistent lower F_0 of the stem-final accent, it should be differentiated from the other accents, which in turn leads us to categorize it as a different group separate from the other accents. This statistical analysis, therefore, confirms the previous speculation that the final accent in stems is less marked than the other types, as predicted by S.-H. Kim's SBG analysis.

(11) The comparison of F_0 (Hz) among the three items within each subject

Subject	Stem-final H		Stem-final H with enclitics		Stem-penult H
S1	112	<	127	<	130
S2	133	<	160	=	160
S3	144	<	155	<	157
S4	125	<	142	<	148
S5	121	<	133	>	132
S6	119	<	128	<	131
S7	96	<	110	=	110
S8	143	<	154	<	156
S9	114	<	126	>	125
S10	128	<	138	<	140

3.2. Experiment 2: the effect of utterance boundary tone

Although experiment 1 supports S.-H. Kim's SBG analysis, it is possible that final high tones are lowered generally in NKK. That is, perhaps there is no difference in the original pitch targets for the tones, but the lower final tone is due to the influence of a final L% boundary tone, as illustrated in (12).

(12) The lowering effect by utterance boundary tone (L%)

a. H_1 [x x x] L%

b. H_2 [x x x x x] L% $\Longrightarrow$ $H_1 < H_2$

Since the high tone H_1 is located on the final syllable of the final accented stem in (12a), a L% boundary tone might affect this high tone, lowering it somewhat. Contrary to H_1, H_2 in (12b) is not adjacent to the L% boundary tone, which then might prevent H_2 from being lowered.

3.2.1. Methods

Noticing that the potential lowering effect in (12) is mediated by the number of syllables, we tested this possibility by comparing the pitch value of the double tones in two environments: word-finally (i.e., in disyllabic words) and non-word-finally (i.e., in trisyllabic words). While the doubly linked high tone in a disyllabic word is adjacent to the following boundary tone (L%), the one in a trisyllabic is not adjacent to the boundary tone (L%) due to the final syllable unlinked to H. Therefore twelve pairs of disyllabic and trisyllabic words were constructed, (13). If there is a general L% boundary effect, the tone in disyllabic words should be lower than that in the matched trisyllabic words.

(13)

Disyllabic words		Trisyllabic words	
kulim	'picture'	kulimca	'shadow'
yein	'woman'	yeinswuk	'inn'
saep	'business'	saepka	'business man'
wisen	'hypocrisy'	wisenca	'hypocrite'
thaca	'typing'	thacaki	'typewriter'
cangnan	'play'	cangnankam	'toy'
kohyang	'hometown'	kohyangcip	'house in hometown'
cenhwa	'telephone'	cenhwaki	'telephone'
yensel	'public speech'	yenselmwun	'script for speech'
kiswul	'technic'	kiswulca	'engineer'
chokcin	'acceleration'	chokcincey	'palpation drug'
pangsa	'radiation'	pangsanung	'radioactivity'

3.2.2. Procedure

The procedure of this experiment follows that of experiment 1. The same speakers who participated in the experiment 1 were asked to read the test words in (13). Two repetitions of each word were collected for a total of the 48 utterances (24 stimuli x 2 repetitions) for each speaker.

3.2.3. Results

For each subject, the mean values for the two types of the stimuli were obtained and used for the analysis, as in (14).

(14)

Type	Number	Mean F_0	Standard deviation
Disyllabic words	240	129.2 Hz	15.5 Hz
Trisyllabic words	240	131.0 Hz	16.0 Hz

The mean F_0 for disyllabic words (129.2 Hz) is slightly lower than for trisyllabic words (131.0 Hz), but this small difference (1.8 Hz) is not statistically significant, $t(238) = -0.886$, $p = 0.3766$. The observed power for this effect size is only 0.1188, so with this number of subjects we only have a 12% chance of statistically detecting a difference of 1.8 Hz at an alpha level of 0.05. However, the smallest detectable difference for a power of 0.95 given the observed measurements is 8.5 Hz. Therefore, we can be 95% confident that the true difference between the means is less than 8.5 Hz. This is much less than the 14 Hz difference observed in experiment 1, and therefore we conclude that final lowering is not the cause of the difference observed in experiment 1.

Looking at the individual subjects we also fail to find a consistent direction of difference between the types. Disyllabic words have a lower mean F_0 for seven subjects, but three subjects show the opposite pattern, (15).

(15) The comparison of F_0 (Hz) between the two items within a subject

Subject	Disyllabic words		Trisyllabic words
S1	116	<	117
S2	136	<	138
S3	142	<	146
S4	143	>	142
S5	124	>	123
S6	119	>	118
S7	103	<	106
S8	153	<	157
S9	121	<	126
S10	134	<	138

Therefore, we cannot reject the null hypothesis that the means are the same, and the substantial difference in effect sizes (14 Hz in experiment 1, 1.8 Hz in experiment 2) argues strongly against final lowering as an explanation for the results in experiment 1. We should note, however, that Chilin Shih (personal communication) has suggested to us that doubled tones may be universally immune from lowering effects; should that be the case then the markedness and lowering explanations would be confounded in NKK and could not be resolved by empirical tests.

4. Discussion and Conclusion

We have considered two accounts of NKK tone that differ in the markedness assigned to final and penult tones in the analyses. N.-J. Kim's OT account proposes that the penult accent is unmarked, whereas S.-H. Kim's SBG account proposes that final accent is unmarked. Assuming that phonological marked elements are usually phonetically implemented with more extreme values, the two accounts then make contrasting predictions about the relative pitch values for tones in different positions. In line with S.-H. Kim's SBG analysis, the results obtained in the first experiment revealed that F_0 peaks in the final accented stems are significantly lower than those in both the related encliticized words and the non-final accented stems. In the second experiment, the finding that there is little pitch difference of the tones in the disyllabic and trisyllabic words eliminates the possibility that the lower final tone is due to the influence of a final L% boundary tone. Therefore, we conclude that only S.-H. Kim's analysis accounts for the significant difference in tone observed in the first experiment and the final accented stems are better analyzed as unaccented stems than as accented stems.

Having established the superiority of S.-H. Kim's analysis, we wish to elaborate some points. First of all, the greater prominence of marked tones in S.-H. Kim's analysis is not due to a direct phonetic reflection of contrasting inputs *per se*, but due rather to a pertinacious difference in metrical structure. The marked tones, lexically marked with right parentheses, are found at the surface in closed feet, whereas the unmarked tones are found in open feet. With the relative strength of closed feet, the only prediction based on SBG Theory is that the tone marked with lexically stored parenthesis is higher than the unmarked tone. Unlike SBG Theory, OT can change the marked tones in the input if necessary, which means we cannot make a coherent prediction within OT.

Also relevant in the discussion of markedness is the range of variation exhibited in different constructions. As we already observed in (4), the high tone of the final accented stems remains in its position when the stem is combined with a vowel-initial enclitic, whereas a high tone shifts to an enclitic when a consonant-initial enclitic is attached to the stem. These two different behaviors of the stem-final tone cannot be found in the stem-penult tone in (2). Since S.-H. Kim's analysis considers stem-final tone unmarked and stem-penult tone marked, it is the unmarked element which shows more behaviors. This is, we believe, a highly desirable situation in that it is natural that less information in the underlying representation leads to more variation in the surface.

Furthermore, if we combine the results in the first and second experiments, we find, using the Scheffé post-hoc repeated measures multiple comparison test, that the double tone is significantly different from both the final and penult tones, giving a third, intermediate level. The summary statistics where the four items are compared pair-wise are shown in (16).

(16) Summary statistics for the Scheffé test on F_0 across the four items

Compared Items	Mean Difference	P-Value
stem-final tone *vs.* double tone	6.7 Hz	0.005
stem-final tone *vs.* stem-final tone + enclitic	13.9 Hz	< 0.0001
stem-final tone *vs.* stem-penult tone	15.5 Hz	< 0.0001
double tone *vs.*. stem-final tone + enclitic	7.2 Hz	0.002
double tone *vs.* stem-penult tone	8.8 Hz	< 0.0001
stem-final tone + enclitic *vs.* stem-penult tone	1.6 Hz	0.917

In this display, the stem-final tone is significantly different from the other items, as shown in the shaded blocks of the first, the second, and the third rows. In the next two rows, the double tones are also significantly different from the other items such as the tones in encliticized words and penult accented stems. As expected, the mean difference is not significant only when stem-final tone in the encliticized words and the stem-penult tone are compared ($p = 0.917$). Given this result, the four types of the stimuli can be divided into three categories, as illustrated in (17).

(17) Categorization of the stimuli in terms of the F_0 value of H

final accented stem	<	double accented stem	<	encliticized word penult accented stem
	≈ 7 Hz		≈ 7 Hz	

It is important to note that the pitch value of the double accented words forms an intermediate category: it is greater than that of the final accented stems by around 7 Hz and less than that of the encliticized words or the penult accented stems by the same amount (around 7 Hz). This observation is consistent with the previous finding that the F_0 in words of the double tone type is relatively lower than that in words of the non-final tone type (Chung 1980:102-107).

Our conjecture about the lower pitch in double tone is that all of the high tones are not necessarily different from one another but that perhaps there is a psychological effect that has to do with the interaction between pitch and time. It can be said that prolonging a somewhat smaller excursion has the same psychological effect as having a very large excursion within a very short amount of time. In other words, since the doubled high tone has a longer pitch excursion, it does not require as much of an increase in the pitch value. In this view, time is traded off against the height of F_0 value. We leave this speculation for further investigation.

References

Chang, Woohyeok. 2002. The use of phonetic evidence to resolve phonological controversies. PhD dissertation, University of Delaware, Newark, DE.

Chung, Kook. 1980. Neutralization in Korean: A functional view. PhD dissertation, University of Texas, Austin, TX.

Chung, Young-Hee. 1991. The lexical tone system of North Kyungsang Korean. PhD dissertation, the Ohio State University, Columbus, OH.

Idsardi, William J. 1994. Open and closed feet in old English. *Linguistic Inquiry* 25: 522-533.

Kenstowicz, Michael, and Hyang-sook Sohn. 1997. Phrasing and focus in North Kyungsang Korean. *MITWPL* 30: 24-48.

Kim, Gyung-Ran. 1988. The pitch-accent system of the Taegu dialect of Korean with emphasis on tone sandhi at the phrasal level. PhD dissertation, University of Hawaii.

Kim, No-Ju. 1993. Default H-tone class in the North Kyungsang Dialect of Korean. *ESCOL* '93 10: 187-198.

Kim, No-Ju. 1997. Tone, segments, and their interaction in North Kyungsang Korean: A correspondence theoretic account. PhD dissertation, The Ohio State University, Columbus, OH.

Kim, Sun-hoi. 1999a. The metrical computation in tone assignment. PhD dissertation, University of Delaware, Newark, DE.

Kim, Sun-hoi. 1999b. Tone, prominence, and simplified footing. *ESCOL '99* 15: 130-141.

William J. Idsardi and Woohyeok Chang
46 East Delaware Avenue

William J. Idsardi and Woohyeok Chang

Department of Linguistics, University of Delaware
Newark, DE 19716-2551

idsardi@udel.edu, cwh@udel.edu

Multiple Questions in Basque*

YOUNGMI JEONG
University of Maryland

1. Background on Multiple Wh-fronting

Extensive research on the syntax of multiple wh-fronting (Rudin 1988, Bošković 1999, 2002, Richards 1997, 2001, Boeckx and Grohmann 2003) has demonstrated the existence of two patterns, the Bulgarian pattern, illustrated in (1), and the Serbo-Croatian pattern, illustrated in (2).

(1) a. Koj kogo kakvo e pital? (Bulgarian)
who whom what is asked
'Who asked whom what?'
b. Koj kakvo kogo e pital?
c. *Kogo kakvo koj e pital?
d. *Kakvo kogo koj e pital?
e. *Kakvo koj kogo e pital?
f. *Kogo koj kakvo e pital?
g. *Koj e pital kogo kakvo
h. *Kogo e pital koj kakvo
i. *Kakvo e pital koj kogo
etc.

As (1) shows, Bulgarian forces the highest wh-word to be the topmost element in the "wh-cluster" (so-called superiority effect), and does not impose any further ordering on the remaining wh-fronting. In Serbo-Croatian, no ordering condition at all is observed:

* Special thanks to Cedric Boeckx, Norbert Hornstein, Howard Lasnik, Juan Uriagereka, and audiences at Berkeley (Berkeley Linguistics Society 29), University of Pennsylvania (Penn Linguistics Colloquium 27), and University of Maryland (Syntax-Semantics Workshop) for important questions and comments.

(2) a. Ko šta gdje kupuje (Serbo-Croatian)
who what where buys
'who buys what where'
b. Ko gdje šta kupuje
c. Šta ko gdje kupuje
d. Šta gdje ko kupuje
e. Gdje ko šta kupuje
f. Gdje šta ko kupuje

However, as Bošković (1997 and subsequent work) has observed, superiority effects surface in Serbo-Croatian as well once we move away from simple, monocalusal contexts. In particular, Bošković has argued that superiority obtains in the language in all contexts where there is evidence for the presence of a complementizer node in overt syntax: matrix questions with overt complementizer, embedded questions, long-distance questions, correlatives, etc. I here illustrate this fact by using long-distance extraction (3).

(3) a. Ko si koga turdio da je istukao?
who are whom claimed that is beaten
'Who did you claim beat whom?'
b. *Koga si ko turdio da je istukao?

2. Previous Accounts

Currently, we have two ways of accounting for the observed patterns: Bošković 1999 and Richards 1997. For Bošković 1999, superiority effects arise as a result of checking a [wh]-feature that is located on C^0. That feature can only be checked once (in his terms, it is an 'Attract-1 feature'), and only by the Closest element, due to Attract Closest, which demands that the Closest matching element be attracted for checking purposes (see Chomsky 1995). Wh-fronting is also induced by the presence of a [Focus]-feature on C^0. That feature, unlike [wh], is an 'Attract-All feature'. Specifically, it has the effect of forcing movement of all matching elements (i.e., wh/focus-phrases). Because of its 'unselective' nature, movement to check [Focus] may take place in any order, as the same number of nodes is ultimately crossed in whatever order of attraction. The combination of [wh] and [Focus] accounts for Bulgarian. For Serbo-Croatian, Bošković claims that in simple contexts C^0 need not be present in overt syntax, hence there is no [wh]-feature active. Unlike [wh], [Focus] can reside on a distinct head, and attract as it does in Bulgarian, that is, in any order. Beyond simple clauses, C^0 (and [wh]) must be present in the language, and the effect of Attract Closest are then being felt.

For Richards, the Bulgarian pattern is not the effect of combining distinct features, but combining distinct principles: Closest Attract (Attract the Closest element; target-perspective), Shortest Move (Move by crossing the minimum number of nodes possible; mover-perspective), and the Principle of Minimal Compliance (PMC), stated in (4).

(4) Principle of Minimal Compliance (PMC)
If a tree contains a dependency headed by H which obeys constraint C, any syntactic object G which H "immediately c-commands" can be ignored for purposes of determining whether C is obeyed by other dependencies.

The effect of Shortest Move forces multiple movement to the same domain to 'tuck-in' (moving to the inner specifier, as opposed to an outer specifier, as moving to the latter would cause movement to cross more nodes than necessary for convergence). The PMC essentially relaxes the order among tucked-in/inner-specifiers. Richards claims that no different account is needed for Serbo-Croatian. The cases where superiority fails to obtain are due to the fact that scrambling has disrupted the 'base' order of wh-elements, so that what gets attracted first is not what is canonically closest.

Both Bošković's and Richards's accounts suffer from problems of their own, but it is not my intention to address these. Rather, I would like to bring new data from Basque that, when correctly interpreted, require a different *kind* of account for multiple wh-fronting. Put simply, either Bošković's or Richard's analysis works as long as languages fall into either the Bulgarian pattern (1) or the Serbo-Croatian patter (2). Basque, which has multiple wh-fronting, as we will see in section (3), doesn't.

3. Some Background on Multiple Wh-questions in Basque

Let us examine the interrogative strategies in Basque.[1] Basque is a predominantly head-final language with free word order. In neutral contexts, the order of arguments is <Nominative/Ergative, Dative, Accusative/Absolutive>.

Basque has several strategies to form multiple questions. The first possibility is to move one wh-phrase and leave the other(s) in situ, as shown in (5). In that strategy is chosen, no superiority effect arises: the fronted wh-phrase need not be the Closest one; and the order of wh-phrases in situ is free.[2]

(5) a. Nork eman dio nori zer?
Who-ERG give aux to whom-DAT what-ABS
'Who gave what to whom?'

[1] For reasons of space, I focus here exclusively on wh-arguments; for general comments on the wh-adjuncts, see Jeong 2003. For an analysis of wh-adjuncts in Basque, see Jeong in progress.

[2] Note, incidentally, that the appearance of *in-situ* wh-phrases to the right of the verb in (5) indicates that the verb in Basque moves leftward, in accordance with Kayne's 1994, 2003 Universal Base (SVO) hypothesis.

b. Nork eman dio zer nori?
Who give aux what to whom
c. Nori eman dio nork zer?
To whom give aux who what
d. Nori eman dio zer nork?
To whom give aux what who
e. Zer eman dio nork nori?
What give aux who to whom
f. Zer eman dio nori nork?
What give aux to whom who

The second option is for Basque to move two wh-phrase and leave the remaining wh-phrase in situ. Here superiority effects between the fronted wh-phrases. Consider (6).

(6) a. Nork nori eman dio zer?
Who-ERG to whom-DAT give aux what-ABS
'Who gave what to whom?'
b. Nork zer eman dio nori?
Who what give aux to whom
c. Nori zer eman dio nork?
To whom what give aux who
d. *Nori nork eman dio zer?
To whom who give aux what
e. *Zer nork eman dio nori?
What who give aux to whom
f. *Zer nori eman dio nork?
What to whom give aux who

However, superiority effects disappear in situations like (6) only if there is a pause/intonational break after the wh-phrase 'violating' superiority. Consider (7). (The pause /intonational break is indicated by #.)

(7) a. Nori # nork eman dio zer?
To whom who give aux what
b. Zer # nork eman dio nori?
What who give aux to whom
c. Zer # nori eman dio nork?
What to whom give aux who

The third, and last strategy for Basque is to move all three wh-phrases. In this case, as originally noted by Ortiz de Urbina (1989), superiority effects arise. However, unlike Bulgarian, Basque imposes a strict ordering among *all* wh-phrases, as we can see in (8).

(8) a. Nork nori zer eman dio?
Who-ERG to whom-DAT what-ABS give aux
'Who gave what to whom?'
b. *Nork zer nori eman dio?
Who what to whom give aux
c. *Nori nork zer eman dio?
To whom who what give aux
d. *Nori zer nork eman dio?
To whom what who give aux
e. *Zer nori nork eman dio?
What to whom who give aux

And just like we saw in (7), superiority effects disappear if there is a pause/intonational break after the wh-phrase 'violating' superiority. Witness (9).

(9) a. Nork zer # nori eman dio?
Who what to whom give aux
b. Nori # nork zer eman dio?
To whom who what give aux
c. Nori zer # nork eman dio?
To whom what who give aux
d. Zer nori # nork eman dio?
What to whom who give aux

Interestingly, whenever superiority fails to obtain ((7) and (9)), replacing the otherwise offending wh-phrase by an aggressively non-D-linked element corresponding to *'wh-the hell'* renders the sentence ungrammatical. Likewise, wh-phrases that remain in situ ((5), (6), (7)) cannot be modified by *'the hell'* (see also Reglero 2003 for a similar observation):

(10) a. Zer # nori eman dio nork?
What to whom give aux who
'who gave what to whom'
b. *Zer arraio nori eman dio nork
what the hell to whom give aux who
'who gave what the hell to whom'
(11) a. Nork erosi du zer?
who-ERG buy AUX what-ABS
'Who bought what?'
b. *Nork erosi du zer arraio?
who-ERG buy AUX what hell-ABS
'Who bought what the hell'

The facts in (10) and (11) are very reminiscent of the fact that D-linked wh-phrases (which cannot be modified by *the hell*, see Pesetsky 1987) are the only type of wh-phrases that may need not front in Bulgarian/Serbo-Croatian, and that may violate superiority in contexts where it otherwise obtains (see Bošković 2002):

(12) Ko je kupio koju knjigu? (Serbo-Croatian)
Who is bought which book
'Who bought which book?'

(13) Koja kniga koj čovek e kupil? (Bulgarian)
Which book which man is bought
'Which man bought which book'

On the basis of this parallelism between Slavic and Basque, I would like to argue that instances of wh-in-situ in Basque are necessarily D-linked (on simplex forms like 'who' and 'what' as covert D-linked phrases like 'which person' and 'which thing' respectively, see Pesetsky 1987). Similarly for wh-phrases 'violating' superiority. Following Grohmann 1998 and Rizzi 2001, among others, I would like to argue that D-linked wh-phrases are 'topics' which target a TopicPhrase that may be located at the left edge of VP (see Rizzi 1997, Belletti 2001, Jayaseelan 2001, among others), or at the left periphery of the clause ('C-domain', see Rizzi 1997). Since topicalization in general does not have to abide by Closest Attract, it is not surprising to find apparent superiority violations with D-linked wh-phrases in Basque.

If the claim just made is correct, instances of multiple wh-fronting with superiority are the only instances of genuine multiple wh-fronting,[3] where fronting takes place to check a [wh] or [focus] feature.[4] From here on, I will set aside instances of topicalization, and propose an analysis of genuine multiple wh-fronting.

4. Analysis

The distinguishing factor between Basque and Bulgarian/Serbo-Croatian is the fact that superiority is pervasive in Basque, while it stops applying after the first instance of fronting in Slavic. Neither Bošković nor Richards predict the Basque pattern; their analyses appear to be dovetail for Slavic. To capture Basque,

[3] The 'the-hell' data, especially (i), removes the possibility that (8) is in effect a situation where all wh-phrases are in situ, in their canonical positions (a possibility pointed out to me by Juan Uriagereka, p.c.), since *arrio* 'the hell' is incompatible with a wh-phrase in situ.

(i) Nork nori zer arraio eman dio?
Who-ERG to whom-DAT what-ABS the-hell give aux
'Who gave what to whom?'

[4] That should not be taken to imply that whenever superiority effects obtain, as in (8), wh-phrases are necessarily non-D-linked. They may be D-linked, and we may be dealing with multiple topics. Although topicalization need not abide by Closest Attract, it need not violate it.

Richards would have to claim that the PMC does not apply in Basque, while it (crucially) does in Slavic. This would be a strange parameter to say the least. For Bošković, Basque poses a different problem altogether. For him, multiple fronting is the result of an [Attract-All] feature ([Focus]). Superiority arises as a result of the presence of an [Attract 1]-feature in C^0 ([wh]). Since Basque shows two instances of superiority, it would require two features with an [Attract-1] quality, alongside the Attract-All [focus]. It's not clear what the additional [Attract-1] feature could be. Moreover, it's not clear how two [Attract-1] features would combine their effects to yield the observed patterns. There should be a hierarchy among the two [Attract-1] features. How this hierarchy could be achieved is by no means obvious.

In light of the shortcomings of the prevailing views on multiple wh-fronting, I would like to develop an alternative. I will focus on Basque first, and return to Slavic at the end of this section.

Following Rizzi 1997 and much subsequent work, I assume that the target of wh-movement is Focus^0. Focus^0 has an Attract-All [Focus] feature, which acts as Bošković claims it does: it is unselective in its pattern of attraction. In addition, Focus^0 has a "V"-type feature forcing verb movement. This is independently needed, as Basque, like so many other languages, requires the finite verb to be adjacent to the focused element(s), a much-discussed phenomenon in Basque linguistics (both traditional and generative) known as 'galdegaia' (see Ortiz de Urbina 1989, 1995, 1999, and Uriagereka 1999, among many others). This is exemplified in (14).

(14) a. Zer egiten duzu zuk hemen?
What do.IMPF AUX you.ERG here
'What are you doing here?'
b. *Zer zuk egiten duzu hemen?

The "V"-feature in Focus^0 must single out the finite verb, not just any verb, since only finite verbs undergo 'galdegaia.' Typically, finite verbs have two features: a T(ense)-feature, and Φ-features. It is the combination of both Φ and T to define a finite verb. In Basque, the finite verb encodes both Ergative/Nominative and object agreement. Importantly, in ditransitives, the dative element triggers agreement. In such situations, the Absolutive/Accusative marker is always restricted to a ø-marker (see Ormazabal and Romero 2002), which I take to mean absence of agreement (as opposed to the more common interpretation of 3^{rd} person agreement in the Basque literature. (The reason for my interpretation will be clear shortly)). Thus, consider (15).

(15) saldu d-ø-izki-o-te
sell D-[øABS]-IZKI-3Dat-3ERG
'(they) sold (it) (to them)

I would like to relate the two Φ-feature exponents on the finite verb to the two features in INFL: Φ and T. I take it that the bundle Φ+T on INFL is related to (i.e., Probes for/Matches) the Ergative/Nominative NP. Infl's T-feature is also related to (Probes for/Matches) *v*, which, following much recent work, I take to be the locus of object (Dative) agreement. So INFL contains information about the Nominative/Ergative NP and the Dative NP. Since Focus^0 matches INFL, it contains that information too, so upon attraction, it attracts the Nominative/Ergative element (related to Φ and T) and the Dative element (related to *v*/T). The Accusative/Absolutive element moves due to the [Attract-All Focus] feature (that ensures overt multiple wh-fronting), as a default, and therefore occupies the last 'slot': <Nom/Erg, Dat, Acc/Abs>. So in effect I am claiming that what forces (iterative) superiority is actually verb movement (and the features it contains), not a 'wh'-feature (unlike Bošković).

My analysis receives striking confirmation from both Basque data not yet discussed, and from Slavic, as I now show.

Consider (16) (taken from Etxepare and Ortiz de Urbina 2002).

(16) a. Nori nork zer iruditu zaio esan duela?
who.to who.ERG what seem AUX say AUX.COMP
'To whom does it seem that who said what?'
b. Nork zer nori iruditu zaio esan duela?

(16) shows that superiority effects are absent in Basque in cases of multiple fronting of wh-phrases originating from different clauses. This is expected under my account since the verb immediately following the wh-phrases signaling the Focus^0-attracting head no longer encodes agreement for all of the wh-phrases (the verb only encodes agreement for a clause-mate wh-phrase).

Let's now turn to Slavic. Independent research on V-movement in Slavic (see Bošković 2001 for review) has established the fact that in Bulgarian, the verb moves to Focus^0 (like Basque, and English). But unlike Basque, the Bulgarian finite verb only encodes one instance agreement (17), not two, so we predict that superiority effects won't be iterative in the language.[5]

[5] The agreement encoded is for the subject, but since I don't assume that feature values enter into the computation of locality (see Boeckx and Jeong 2002 for extensive discussion), I don't predict that superiority in Bulgarian only affects subject wh-phrases. Rather, I predict that it only affects agreeing wh-phrases. Thus, I predict, correctly, that Focus^0 will attract the dative wh-phrase over the accusative wh-phrase in the absence of a nominative wh-phrase (i). Likewise, I don't predict that non-wh-subjects will raise to Focus^{0}in Bulgarian since they don't match the [Focus] feature. For fuller discussion, see Jeong (in progress).

(i) kogo kakvo/*kakvo kogo e pital Ivan
whom what/ what whom is asked Ivan
'What did Ivan ask to whom?'

(17) (Petko) mi go dade
Petko me.Dat it.Acc gave.3sg
'Petko gave it to me'

As for Serbo-Croatian (or Russian), it has been shown that the language lacks V-movement to Focus^0 (with Bošković, I assume that in such cases Focus^0 is not present at all in overt syntax). We then predict no superiority effect.[6]

All in all, our account of superiority based on agreement and verb movement to the [wh]-attracting head captures all the attested patterns of multiple wh-fronting, not only in Basque, but also in Slavic. It is therefore superior to Bošković's and Richards's accounts.

5. Conclusion

In this paper I have investigated a special feature of multiple wh-questions in Basque, the existence of superiority with *each* instance of wh-fronting, and show how it forces us to reconsider our understanding of the mechanisms of multiple wh-fronting in general, since standard accounts based on Slavic (Bošković 1999, Richards 1997) predict a relaxation of superiority after the first instance of wh-fronting. As an alternative I have proposed that the role of the verb is crucial in determining the specific patterns of multiple wh-fronting found across languages. If correct, my analysis shows that head-movement cannot be entirely shifted from narrow syntax into the PF-component, as argued by Chomsky and others in recent work, as head-dependencies are the key factors in the syntactic organization of multiple wh-fronting. It also shows that because it depends on such uninterpretable features as Φ-features on finite verbs, superiority is a narrow-syntax requirement, and not an interpretive effect (as argued by Chierchia 1991 and Hornstein 1995).

References

Belletti, Adriana. 2001. Aspects of the low IP area. Ms., University of Siena.

Boeckx, Cedric and Kleanthes K. Grohmann (eds). 2003. *Multiple Wh-fronting*. Amsterdam: John Benjamins.

Boeckx, Cedric and Youngmi Jeong. 2002. The fine structure of syntactic intervention. Ms., University of Maryland.

Bošković, Željko. 1997. Superiority effects with multiple wh-fronting in Serbo-Croatian. *Lingua* 102, 1-20.

[6] When superiority emerges in Serbo-Croatian, Focus^0/C^0 is overt. Although V-movement does not take place, I assume that there is a featural relation between Focus^0/C^0 and the finite verb (see Pesetsky and Torrego 2001), which is enough to force superiority. Again, since the finite verb in Serbo-Croatian only encodes one instance of agreement, we predict that superiority will not be iterative.

Bošković, Željko. 1999. On multiple feature checking: Multiple wh-fronting and multiple head-movement. In *Working Minimalism*, S. Epstein and N. Hornstein (eds), 159-187. Cambridge, Mass.: MIT Press.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface*. London: Elsevier.

Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33, 351-383.

Chierchia, Gennaro. 1991. Functional WH and weak crossover. In *Proceedings of WCCFL 10*, 44-55. Stanford, Calif.: CLSI.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Etxepare, Ricardo and Jon Ortiz de Urbina. 2002. Focalization. Ms., University of Deusto.

Grohmann, Kleanthes K. 1998. Syntactic inquiries into discourse restrictions on multiple questions. *Groninger Arbeiten zur Germanisitschen Linguistik* 42, 1-60.

Hornstein, Norbert. 1995. *Logical Form*. Oxford: Blackwell.

Jayaseelan, K.A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55, 39-75.

Jeong, Youngmi. 2003. Deriving anti-superiority effects. In *Multiple Wh-fronting*, ed. C. Boeckx and K.K. Grohmann. Amsterdam: John Benjamins.

Jeong, Youngmi. In progress. Multiple wh-fronting in Basque: causes and consequences. Ms., University of Maryland.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 2003. Antisymmetry and Japanese. Ms., NYU.

Ormazabal, Javier, and Juan. Romero. 2002. Agreement restrictions. Ms., Un iversity of the Basque Country and University of Alcala. [To appear in *Natural Language and Linguistic Theory*.]

Ortiz de Urbina, Jon. 1989. *Parameters in the Grammar of Basque*. Dordrecht: Foris.

Ortiz de Urbina, Jon. 1995. "Residual verb second and verb first in Basque". In *Discourse Configurational Languages*, K. É. Kiss (ed), 99-121. New York and Oxford: Oxford University Press.

Ortiz de Urbina, Jon. 1999. Focus in Basque. In *The Grammar of Focus*, G. Rebuschi and L. Tuller (eds), 311-333. Amsterdam and Philadelphia: John Benjamins.

Pesetsky, David. 1987. Wh-in situ: Movement an unselective binding. In *The Representation of (In)definitiness*, E. J. Reuland and A. G. B. ter Meulen (eds), 98-129. Cambridge, Mass.: MIT Press.

Pesetsky, David and Esther Torrego. 2001. T-to-C: causes and consequences. In *Ken Hale: A life in language*, ed. M. Kensowicz, 355-426. Cambridge, Mass.: MIT Press.

Reglero, Lara. 2003. Non-wh-fronting in Basque. In *Multiple Wh-fronting*, ed. C. Boeckx and K.K. Grohmann. Amsterdam: John Benjamins.

Richards, Norvin. 1997. What Moves Where When in Which Language. PhD dissertation, MIT, Cambridge, Mass.

Richards, Norvin. 2001. *Movement in language*. Oxford: Oxford University Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. L. Haegeman, 281-337. Dordrecht: Kluwer.

Rizzi, Luigi. 2001. Reconstruction, weak island sensitivity, and agreement. In *Semantic interfaces*, ed. C. Cecchetto, G. Chierchia, and M.T. Guasti, 145-176. Stanford: CSLI.

Rudin, Catherine. 1988. "On multiple questions and multiple *wh*-fronting". *Natural Language and Linguistic Theory* 6: 445-501.

Uriagereka, Juan. 1999. Minimal restrictions on Basque movement. *Natural Language and Linguistic Theory* 17, 403-444.

Dept. of Linguistics
University of Maryland
College Park, MD 20742

yjeong@wam.umd.edu

Nominal and Adjectival Roots in North Saami Verbs*

MARIT JULIEN
University of Tromsø

0. Introduction

According to recent proposals such as Marantz (1997, 2001) and Borer (2000) roots have no lexical category. The categories Noun, Verb, and Adjective result from combining a categoryless root with functional projections of nominal, verbal or adjectival nature, respectively.

While this appears to give a neat explanation of languages like English, for example, where many roots can function as nouns, as verbs, or as adjectives, the categories only differentiated by the inflection, I will nevertheless claim that even before they are inserted into syntactic structures, roots fall into different classes that correspond to the lexical categories noun, verb, and adjective. My claim is based on some observations of the semantic properties of derived verbs in North Saami. In this language, the meaning of stative, inchoative and causative verbs built on adjectival roots contrasts systematically with the meaning of verbs built on nominal roots. The meaning contrast must ultimately stem from the roots. That is, the class membership of the root affects the computation of the meaning of the derived verb.

The paper is organized as follows. In section 1 I present the basic facts about deadjectival and denominal verbs in North Saami. In section 2 I put forward a proposal concerning the syntactic representation of these verbs, and I show how the observed meaning contrast can be explained within the proposed framework. In section 3 I address some apparent exceptions to my analysis, and in section 4, I add some more comments on Marantz (1997). My conclusions are summed up in section 6.

1. Deadjectival and denominal verbs in North Saami

Let us begin by looking at the North Saami stative verbs in (1), which arguably are based on adjectival roots. Note that the citation forms of the adjectives include a marker of nominative singular, and that the verbs are given in their infinitival form, which includes the infinitival marker *–t* (after vowels) or *–it* (after consonants). The derivational marker is the suffix between the root and the infinitival marker. (The monophthongization /ie/>/i/ in (1e) is one instance of a more general phenomenon in the language.)

---

* The research for this paper was financially supported by the Norwegian Research Council, grant no. 141687/540. In developing the ideas presented here, I have received helpful input from Berit Anne Bals, Harald Gaski, and Peter Svenonius. Ollu giitu! I also thank the audiences at Grammar in Focus 2003, Lund, BLS 29, Berkeley, and the Thursday Night Linguistics Seminar, University of Tromsø, for their feedback.

(1) ADJECTIVE			STATIVE VERB	
a. ruoks-at	'red'	>	ruoks-á-t	'appear red'
b. bahč-a	'bitter'	>	bahč-ist-it	'smell or taste bad'
c. bahč-a	'bitter'	>	bahč-iid-it	'smell or taste bad'
d. liekk-as	'warm'	>	liekk-ist-it	'feel warm'
e. liekk-as	'warm'	>	likk-iid-it	'feel warm (by direct contact)'

In North Saami, stative verbs can also be formed from nominal roots. Some examples are given in (2). (In (2a) and (2b) we see another pervasive feature of North Saami morphophonology, namely, stem consonant gradation, as exemplified by the alternations *tnj* ~ *dnj* and *tn* ~ *dn*.)

(2) NOUN			STATIVE VERB	
a. botnji	'twist'	>	bodnj-á-t	'be twisted, bent'
b. ritni	'rime'	>	ridn-á-t	'be covered with rime'
c. ballu	'fear'	>	ball-a-t	'fear, be afraid'
d. bárti	'trouble'	>	bárti-d-it	'be in trouble'
e. illu	'joy'	>	illu-d-it	'be happy'

Now note, firstly, that in (1) and (2) the choice of verbalizing suffix is idiosyncratic for each root. Further, there is a systematic meaning difference between the verbs in the two groups. While the verbs in (1) essentially mean 'be Adjective', the verbs in (2) do *not* mean 'be Noun'; they mean roughly 'have Noun'. However, there is no obvious reason why the derived verbs in (2) should not mean 'be Noun'. From a pragmatic point of view, we would expect a verb meaning 'be joy', for example, to be perfectly acceptable. Moreover, the nominals in (2) may well combine with the copula *leat* 'be' to give expressions like *leat ritni* 'be rime', *leat illu* 'be joy'. Hence, the source of the meanings that we see in (2) must be sought in the grammar, or more precisely, in the syntactic and semantic relations holding between the elements that the complex verbs are made up of.

From adjectival roots it is also possible to derive inchoative verbs in North Saami. The examples in (3) illustrate this. Other inchoative verbs are formed from nominal roots, as in the examples in (4).

(3) ADJECTIVE			INCHOATIVE VERB	
a. dimis	'soft'	>	dipm-a-t	'become soft'
b. goikkis	'dry'	>	goik-a-t	'become dry'
c. obba	'cloudy'	>	obb-e-t	'become cloudy'
d. guhkki	'long'	>	guhkk-u-t	'become long(er)'
e. lossat	'heavy'	>	loss-u-t	'become heavy/heavier'
f. stuoris	'big'	>	stuorr-u-t	'become big(ger)'
g. bahča	'bitter'	>	bahča-g-it	'become bitter, get a bad taste'
h. ruoksat	'red'	>	ruvss-od-it	'turn red, blush'
i. seavdnjat	'dark'	>	sevnnj-od-it	'become dark'

(4) NOUN			INCHOATIVE VERB	
a. čáhci	'water'	>	čáhc-u-t	'become wet'
b. ruobbi	'rash'	>	ruobb-u-t	'get a rash'
c. bárti	'trouble'	>	bárttá-sk-it	'get into trouble'
d. vašši	'hatred'	>	vaššá-sk-it	'get furiously angry'
e. dállu	'farm'	>	dálui-duvva-t	'get a farm'

f.	idja	'night'	>	idja-duvva-t	'be overtaken by night'
g.	vealgi	'debt'	>	vealggái-duvva-t	'get (deeply) into debt'
h.	bihkka	'tar'	>	bihka-huvva-t	'get smeared with tar'
i.	dálvi	'winter'	>	dálvvá-huvva-t	'be overtaken by winter'
j.	sálti	'salt'	>	sáltá-šuvva-t	'absorb salt; become salt enough'

Strikingly, we find a meaning contrast between (3) on the one hand and (4) on the other that is parallel to the contrast between (1) and (2). While the verbs in (3) mean 'become Adjective', the verbs in (4) do not mean 'become Noun', they mean 'get to have Noun' or 'be affected by Noun'.

Even more strikingly, if we derive *causative* verbs from adjectival and nominal roots, respectively, as in (5) and (6), we get a similar effect once more. Causative verbs formed from adjectival roots mean 'cause to be/become Adjective', while causative verbs formed from nominal roots mean 'cause to have/get Noun'.

(5)	ADJECTIVE			CAUSATIVE VERB	
a.	buorre	'good'	>	buori-d-it	'improve$_{\text{TRANS}}$'
b.	čuorbi	'clumsy'	>	čuorbbi-d-it	'make clumsy'
c.	guhkki	'long'	>	guhki-d-it	'lengthen'
d.	hálbi	'cheap'	>	hálbbi-d-it	'make cheap(er)'
e.	liekkas	'warm'	>	liekka-d-it	'warm up, heat'
f.	stuoris	'big'	>	stuori-d-it	'enlarge'

(6)	NOUN			CAUSATIVE VERB	
a.	bihkka	'tar'	>	bihkka-d-it	'cover with tar'
b.	ruhta	'money'	>	ruhta-d-it	'finance'
c.	vearju	'weapon'	>	vearju-d-it	'arm'

The generalization is that in deadjectival verbs, the root is ultimately predicated of the lower argument (that is, of the only argument if the verb is monovalent), while this is not the case with denominal verbs.

2. The source of the meaning contrast

The question is now: What is the source of this meaning contrast? A priori, it could, for example, be the verbalizing suffixes that are responsible. However, for one thing, in some cases one and the same verbalizer can appear with adjectival or nominal roots. One example is the *-á-* that we see in (1) and (2), another example is the *-u-* that we find in (3) and (4), and a third example is the *-d-* that appears in all the verbs in (5) and (6). This fact poses a problem for such an explanation.

For another, if the source of the contrast is the verbalizers we would have to explain why some roots combine only with one type of verbalizer while other roots only combine with another type of verbalizer. Moreover, a similar meaning contrast is seen in English zero causative derivations, as (7) and (8) illustrate.

(7)a. clear (a screen) 'make clear'
b. dry (the clothes) 'make dry'

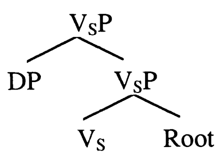
(8)a. saddle (a horse) 'attach a saddle to'
b. man (the torpedoes) 'put men to'

Hence, I think we have to conclude that the meaning contrast is ultimately a consequence of the roots being of different categories. In fact, even Marantz (2001) acknowledges that roots are of different semantic types. Among other things, he observes that some roots denote entities while others denote states. Although this is a relatively trivial observation, it is crucial in the present context. The distinction between entities and states corresponds to the distinction between nouns and adjectives, and it seems clear to me that it lies at the heart of the meaning contrast that we have seen between classes of derived verbs.

To show this, I will present an analysis that is based on what might be called a syntactic decomposition approach to the verbs in question. In an approach of this type it is assumed that if a verb denotes a semantically complex event, each part of the event is represented in the syntax, and that alternations with respect to event structure and argument structure are tightly connected to syntactic alternations. Works such as Hale & Keyser (1993, 1998, 2002), Arad (1998), Travis (1992, 2000), Pylkkänen (2002), and Ramchand (2002) are relevant examples.

On the syntactic decomposition approach, the stative verbs in (1) and (2) must minimally involve a VP with the syntactic structure shown in (9). Here the root has combined with a verbal element, designated as V_S (where S stands for *stative*), which projects a V_SP and thereby accommodates in its Spec the subject of the state, that is, the argument that the state holds of. The verbalizing suffixes in (1) and (2) are then realizations of this V_S element. On top of V_SP there will be inflectional elements, but I will ignore this here and in the following, since the focus is on the formation of the verb itself.

(9)

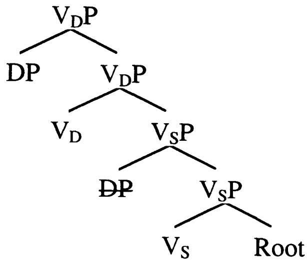


Concerning the idiosyncratic combinations of roots and verbalizers that we saw in (1) and (2) (and also in (3) and (4)), such idiosyncratic combinations are traditionally seen as being typical of so-called lexical derivation. However, Marantz (1997) argues that the properties associated with 'lexical' word formation are restricted to a specific syntactic domain—namely, the domain below the syntactic head that introduces the agent. The verbs in (1) and (2) do not involve an agent. Hence, on Marantz' approach, the observed idiosyncrasies are not unexpected.

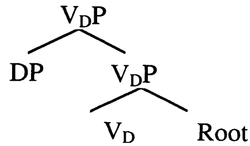
The inchoative verbs in (3) and (4) are semantically more complex than stative verbs. While stative verbs describe situations that have no internal temporal structure, inchoative verbs describe events that consist of a processual part and possibly a result state. If verbal meaning is built compositionally in the syntax, as I am assuming here, it follows that inchoative verbs involve a verbal head that encodes the process, as well as, if the verb is telic, a head encoding the result state. That is, a *telic* inchoative verb involves a VP with the syntactic structure shown in (10a), whereas an *atelic* inchoative verb involves a VP with the syntactic structure shown in (10b).

In (10a), a processual verbal head V_D (D for 'dynamic') is merged over the projection of the stative V_S. The argument DP that starts out in Spec-V_SP, where it is interpreted as the subject of the state, raises to Spec-V_DP and gets the additional interpretation that it is also the undergoer of the process (see Ramchand 2002). In (10b), the V_S head is missing, and consequently, the root combines directly with the V_D head, and the argument of the verb, which is located in Spec-V_DP, is only interpreted as the undergoer of the process.

(10)a.



b.



Derived inchoative verbs in North Saami can be telic or atelic. This is seen in (3f), for example: *stuorrut* can mean both 'get big' and 'get bigger'. If it means 'get big', the result state is defined, but not if it means 'get bigger'. However, the verbs formed with the suffixes *–(i)duvva–*, *–huvva–*, and *–šuvva–*, as in (4e)–(4j), are only telic. As a consequence, while *stuorrut* is compatible with a degree adverbial like *veahá* 'a little' —see (11a), *vealggáiduvvat* 'get into debt' is not—see (11b).

(11)a.

Biret	lea	sturro-n	veahá	dan rájes go	mun	oidnen	su.
Biret.NOM	is	get.bigger-PTC	a little	since	I	saw	her

'Biret has become a little bigger since I saw her.'

b.

Biret	lea	vealggáiduvva-n	(*veahá).
Biret.NOM	is	get.into.debt-PTC	a little

'Biret has gotten (*a little) into debt .'

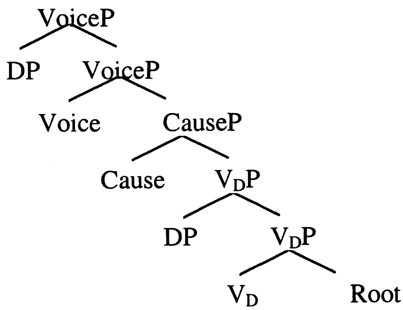
And interestingly, the suffixes *–(i)duvva–*, *–huvva–*, and *–šuvva–* appear to be morphologically complex. The element *–d–*, for example, appears on its own as a verbalizer in (2d)–(2e), where I analyzed it as a realization of V_S. I therefore take *–d–* to represent V_S also in (4e), (4f), and (4g), and *–h–* and *–š–* in (4h), (4i), and (4j) to be like *–d–*. The *–uvva–* part of the suffix is then a realization of the processual head V_D in these examples. It now follows that the modifier *veahá* is felicitous with *stuorrut* but not with *vealggáiduvvat*. Pragmatically, it is of course possible to get more or less deeply into debt. But grammatically, whenever *vealggáiduvvat* is true of a subject, the resulting state also necessarily holds of that subject. It cannot then hold just a little. Because of this, modifiers like *veahá* are not appropriate. *Stuorrut*, on the other hand, does not necessarily encode a result state, and accordingly, it can be modified by degree adverbials.

As for the derived causative verbs in (5) and (6), I take the Cause head to be separate from Voice, the element that introduces the external argument. The reason is that when causatives are passivized, the causative marker is retained, and the passive marker *–uvvo–* appears outside it, as demonstrated in (12). If the passive marker is the realization of a [–active] Voice head, it follows that the causativizer is separate from and below the Voice head (cf. Pylkkänen 2002).

(12) ACTIVE			PASSIVE	
a.	buori-d-it	'improve$_{\text{TRANS}}$' (5a)	> buori-d-uvvo-t	'be improved'
b.	guhki-d-it	'lengthen' (5c)	> guhki-d-uvvo-t	'be lengthened'
c.	hálbbi-d-it	'make cheap(er)' (5d)	> hálbbi-d-uvvo-t	'be made cheap(er)'
d.	ruhta-d-it	'finance' (6b)	> ruhta-d-uvvo-t	'be financed'

Hence, the VP domains of causative verbs minimally have the syntactic structure shown in (13). Note that In addition, at least some of the derived causative verbs also probably include a V_S projection, which is omitted in (13).

(12)



Concerning the claim that roots with substantial content can be of different semantic types, I suggested above that nominal roots denote entities, while adjectival roots denote states. This needs some modification: nominal roots do not exactly denote entities. But they are arguably not predicates either. For example, Zamparelli (2000) argues that a nominal predicate contains various functional projections in addition to the root. Embedded in the predicate, there is a projection that denote the kind, that is, the whole class of entities associated with the meaning of the nominal root, possibly modified by adjectives. I take reference to entities to be the result of determiners and quantifiers having operated on, that is, quantified and possibly specified, the reference of the kind-referring projection (see, e.g., Longobardi 2001). This means that nominal roots can be characterised as entity-denoting in the sense that they denote classes of entities, although they do not denote individual entities, since their denotations are not specified or quantified. It follows that nominal roots are of the same semantic type as full DPs. An indication that this is not too far from the truth is found in Scandinavian. From the Norwegian examples in (14) we see that a bare singular noun can be an argument just as well as a singular noun with a determiner (see Borthen 1999, 2000 for a more detailed treatment). The main difference is that the bare noun does not give any indication of grammatical number.

(14)a. Det vart sett **ein** **ulv** her i går.
it became seen a wolf here yesterday
'A wolf was seen here yesterday.'

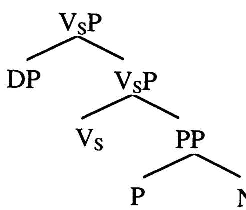
b. Det vart sett **ulv** her i går.
it became seen wolf here yesterday
'One or more wolves were seen here yesterday.'

This means that nominal roots do not exactly denote entities; they denote kinds, that is, classes of entities. They refer to entities when their reference is quantified and possibly specified by functional elements such as determiners and quantifiers (cf. Zamparelli 2000, Longobardi 2001). I nevertheless conclude that determiners are not type-changing—what they do is operate on the reference of the nominal in various ways. I will therefore use the term 'entity-denoting' to characterize nominal roots in the following.

Now let us consider the derivation of stative verbs. I assume that V_S requires a state-denoting element as its complement, and that V_S together with its complement also denotes a state. Thus, when the complement is an adjectival root, there are no problems. However, if the complement is a nominal root, the result is ungrammatical and has no interpretation, because the nominal root is of the wrong semantic type.

One way to save the derivation is to merge the nominal root with a preposition before it combines with the verbalizer. This means that what I showed in (9) is actually the VP-structure of *deadjectival* stative verbs. *Denominal* stative verbs have the VP-structure shown in (15). Here the complement of V_S is a PP, which is of the right semantic type, and a state reading results. But because of the presence of the preposition, the nominal root is not directly predicated of the subject of the state. That is, instead of the meaning 'be Root', we get 'be P Root', which can be interpreted as 'have Root', since HAVE is the result of combining BE and P (as proposed by Kayne 1993). Otherwise, the exact relation that the P represents will be determined in a pragmatic fashion (see, for example, the verbs in (4)).

(15)



Exactly the same reasoning applies to denominal inchoative and causative verbs. The verbalizer requires a state-denoting complement. Because of this, it can combine directly with an adjectival root. Nominal roots, on the other hand, must combine with a preposition first, and the reading we get is 'become P Noun' or 'get to have Noun' for inchoatives, and 'cause to be P Root' or 'cause to have Root' for causatives.

That denominal verbs involve an abstract preposition was originally proposed by Hale & Keyser (1993), mainly on the basis of English verbs like *saddle*, whose meaning appears to involve a prepositional element (see (8a)). Now while this preposition is never visible in English, it sometimes *is* visible in North Saami.[1] Consider again the denominal verbs *dáluiduvvat* in (4e) and *vealggáiduvvat* in (4g), formed from the nouns *dállu* 'farm' and *vealgi* 'debt'. The *–i–* that follows the root in these verbs is not a part of the root, and arguably not a part of the verbalizer either. In fact, it is identical to the marker of illative case, which can be

[1] As we have seen, the verbalizers also have their own phonological realization in North Saami. This means that denominal and deadjectival verbs in North Saami are formed by incorporation and not by conflation, in the terms of Hale & Keyser (2002).

analyzed as a reflex of a preposition. The regular illative singular forms of *dállu* and *vealgi* are shown in (16). (Note that the final *–i* in *vealgi* becomes an *–á* in front of the illative marker. This is a regular alternation which also occurs in *vealggáiduvvat*.)

	NOMINATIVE SG			ILLATIVE SG	
(16) a.	dállu	'farm'	>	dállu**i**	'to a/the farm'
b.	vealgi	'debt'	>	vealgá**i**	'to a/the debt'

We see that the only difference between the forms used in verbalization and the regular illative singular forms is that the stem consonants appear in the weak grade in the former but in the strong grade in latter. I am not sure what to make of the grade contrast, but I nevertheless believe that associating the *–i–* with the P element is not entirely without justification, especially since the *–i–* in question is only found with nominal roots, as Nielsen (1926) points out.

3. Verbs meaning 'become Noun'

At this point, it has to be revealed that North Saami does in fact have a few denominal inchoative verbs meaning 'become Noun'. Some of them are shown in (17).

	NOUN			INCHOATIVE VERB	
(17) a.	dálvi	'winter'	>	dálv-a-t	'turn winter'
b.	geassi	'summer'	>	geass-u-t	'turn summer'
c.	fiertu	'fine weather'	>	firt-e-t	'become fine weather'
d.	gállu	'cold weather'	>	gálu-d-it	'become cold weather'
e.	mosku	'fog in the dark'	>	moskku-d-it	'become foggy in the dark'
f.	duottar	'bare mountain'	>	duottar-duvva-t	'become bare mountain'
g.	šalka	'firm track in snow'	>	šalka-luvva-t	'become firm track in snow'
h.	čahki	'hard lump of snow'	>	čahki-luvva-t	'form into lumps (of snow)'
i.	gea'gi	'stone'	>	gea'g-u-t	'turn into stone'
j.	hilla	'ember'	>	hilla-luvva-t	'turn into embers'
k.	olmmái	'man'	>	olmmái-duvva-t	'become a man'

If my account of the verbs meaning 'become P Noun' is correct, the existence of the verbs in (17) is unexpected. It is striking, though, that the majority of verbs of this type are based on roots that refer to meteorological phenomena, as in the examples (17a)–(17e), or to the terrain, as in (17f), or to the conditions for travelling, as in (17g) and (17h).

As Josefsson (1997) points out, some roots denote concepts that do not clearly belong one or the other ontological category—that is, their semantic class membership is somewhat vague—and consequently, the grammatical behavior of these roots is compatible with more than one semantic type. For the roots that we see in (17a)–(17h), it does not seem far-fetched to claim that they can be conceptualized as entities, which gives rise to the nouns seen in the left-hand column, or as states, which allows them to combine directly with an inchoative verbalizer and yield a verb which will be used to predicate the state of a subject (which in the case of a meteorological verb will be phonologically null.) For example, Nielsen (1932:584) says that *duottarduvvat* (in (17f)) means 'assume the character of bare mountain'. I take this to mean that in this verb, the root *duottar* is seen from a state perspective rather than from an entity perspective.

For the verb in (17k), on the other hand, the analysis I have just sketched may seem less plausible. However, an informant tells me that the precise meaning of *olmmáiduvvat* is 'come to be *like* a man'. Hence, it is arguably the case even here that the root is conceptualized as state-denoting, although it normally would be entity-denoting. It should be noted that very few inchoative denominal verbs in North Saami do not describe meteorological phenomena or other aspects of the natural environment. Thus, there are not many verbs like *olmmáiduvvat* in the language. They cannot be freely formed, in contrast to the denominal verbs involving a P. That is, *olmmáiduvvat* must be listed in the lexicon. From the listed meaning it follows that in this verb, the state-like aspect of the root *olmmái* must be put to the front.

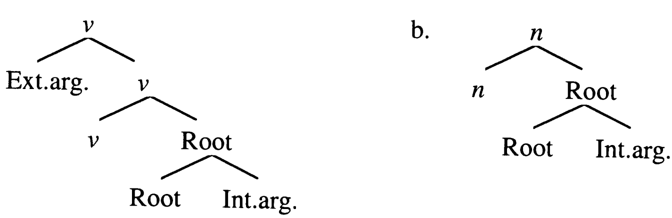
4. Marantz (1997) again

If we now return to Marantz (1997), it seems to me that my claim that roots are of different semantic types is not necessarily in conflict with Marantz' argumentation. Marantz' core observation is that external arguments are not syntactically included in nominalizations, hence the ungrammaticality of (18).

(18) * Mary's growth of tomatoes

The explanation that Marantz offers is that the root has no category, and that when it appears as a transitive verb, it is combined with a *v* head that verbalizes the root and also introduces the external argument in its Spec, as sketched in (19a). (An unaccusative verb, on the other hand, would involve a *v* head that does not allow an external argument to appear in its Spec.) But when the root is inserted in a nominal environment, as in (19b), there is no *v* and consequently no position where an external agent could appear unless the existence of that agent is implied by the root itself. Since *grow* implies internal causation, *growth* does not allow an external agent to be realized.

(19)a. b.

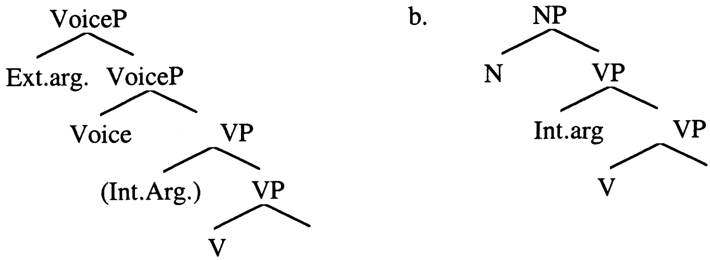


If a root of the *destroy*-type is nominalized, on the other hand, as in (20), the possessor can be interpreted as the agent, since the semantics of the root itself implies the presence of an agent. Thus, the analysis in Marantz (1997) gives the desired result concerning the distribution of external agents in nominalizations.

(20) John's destruction of his room

But as far as I can see, so could an analysis based on the view that roots are specified for category, or at least that the semantic type of the root matters for word formation. After Rivero (1990), Travis (1992), and Kratzer (1996), the idea that the external argument is introduced by a syntactic head distinct from V has become a fairly standard assumption. It follows from this assumption that the agent introducing head, that is Voice, must be above V_D in processual transitive verbs and above V_S in stative transitive verbs. That is, verbs with external arguments involve at least the structure shown in (21a), where V stands for V_S or V_D.

(21)a.



To keep the external argument out of nominalized verbs, all we need to assume in addition is that the nominalizer combines with a constituent that is smaller than VoiceP. In the spirit of Alexiadou (2001) we could say that eventive nominalizations like *destruction* and *growth* contain a V_D projection, and that the nominalizer combines with V_DP. Result state nominals, such as *break*, involve a V_S head. In short, nominalizations minimally have the structure in (20b). In this structure, V can stand for a verbal root, or it can stand for a verbalizer that has a root of some other class as its complement.

5. Conclusions

The preceding discussion has shown, firstly, that derived stative, inchoative, and causative verbs in North Saami can successfully be analyzed within a framework based on syntactic decomposition of complex events. Secondly, the properties of these verbs indicate that roots belong to different semantic classes, corresponding to syntactic categories, and that the semantic class membership of the root can have syntactic and semantic consequences in verbalization.

References

Alexiadou, Artemis. 2001. *Functional structure in nominals.* Amsterdam: John Benjamins.

Arad, Maya. 1998. VP-structure and the syntax-lexicon interface. PhD dissertation, University College London.

Borer, Hagit. 2000. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. Ms., University of Southern California.

Borthen, Kaja. 1999. Bare singulars in Norwegian. *University of Trondheim Working Papers in Linguistics* 32.

Borthen, Kaja. 2000. The distribution pattern of Norwegian bare singulars. In Christina Czinglar, Katharina Köhler, Erica Thrift, Erik Jan van der Torre, & Malte Zimmermann (eds.) *Proceedings of ConSole VIII*, 29–44. Leiden: SOLE.

Hale, Kenneth and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.) *The view from Building 20*, 53–109. Cambridge, MA: MIT Press.

Hale, Kenneth and Samuel Jay Keyser. 1998. The basic elements of argument structure. *MIT Working Papers in Linguistics* 32: 73–118.

Hale, Kenneth and Samuel Jay Keyser. 2002. *Prolegomena to a theory of argument structure*. Cambridge, MA: MIT Press.

Josefsson, Gunlög. 1997. *On the principles of word formation in Swedish*. Lund University Press. [Published 1998 as *Minimal words in a minimal syntax* by John Benjamins, Amsterdam.]

Kayne, Richard S. 1993. Toward a modular theory of auxiliary selection. *Studia Linguistica* 47: 3–31.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.) *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer.

Longobardi, Giuseppe. 2001. How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Natural Language Semantics* 9: 335–369.

Marantz, Alec. 1997. No escape from syntax. *University of Pennsylvania Working Papers in Linguistics* 4(2): 201–225.

Marantz, Alec. 2001. Words. Talk given at the West Coast Conference of Formal Linguistics 20, University of Southern California, Los Angeles, February 2001.

Nielsen. Konrad. 1926. *Lærebok i lappisk (samisk), vol. 1, Grammatikk [Lapp (Saami) primer, vol. 1, Grammar]*. Oslo: Brøgger. Reprinted 1979, Oslo: Universitetsforlaget.

Nielsen. Konrad. 1932. *Lappisk (samisk) ordbok/Lapp dictionary, vol. 1, A–F*. Oslo: Aschehoug. Reprinted 1979, Oslo: Universitetsforlaget.

Pylkkänen, Liina. 2002. Introducing arguments. PhD dissertation, MIT.

Ramchand, Gillian. 2002. Aktionsart, L-syntax and selection. In Henk J. Verkuyl (ed.) *Proceedings of the Perspectives on Aspect Conference*, 1–15. Utrecht: OTS.

Rivero, María-Luisa. 1990. The location of nonactive voice in Albanian and Modern Greek. *Linguistic Inquiry* 21: 135–146.

Travis, Lisa deMena. 1992. Inner Aspect and the structure of VP. *Cahiers de Linguistique de l'UQAM* 1: 130–146.

Travis, Lisa. 2000. Event structure in syntax. In Carol Tenny & James Pustejovsky (eds.) *Events as grammatical objects*, 145–185. Stanford, CA: CSLI Publications.

Zamparelli, Roberto. 2000. *Layers in the determiner phrase*. New York: Garland.

Marit Julien
Faculty of Humanities
University of Tromsø
NO–9037 Tromsø
Norway

Marit.Julien@hum.uit.no

Syllabically Conditioned Perceptual Epenthesis

BARIS KABAK and WILLIAM IDSARDI
University of Delaware

0. Introduction

This paper focuses on perceptual epenthesis; a phenomenon where listeners perceive illusory vowels within consonant clusters which deviate from the phonotactic norms of their native language (see Dupoux et al., 1999). We present results from an experiment on Korean listeners' perception of English consonant clusters which replicates and extends previous studies on Japanese. Our primary aim is to tease apart two explanations for perceptual epenthesis which are confounded in the Japanese studies: consonantal contact violations and syllable structure violations. In light of our results, we suggest here that perceptual epenthesis is caused by syllable structure violations rather than illicit consonantal contact. In addition, we show that speech perception is not always governed by the same system of rules and restrictions that govern speech production. We discuss the consequences of the non-isomorphism between speech production and perception observed in our experiment in the context of the P-map hypothesis (Steriade, 2001a, b). Furthermore, we show that frequency-based analyses fail to account for our results.

1. Background and Problem

The ability to discriminate non-native phonological contrasts deteriorates during the first year of life (see Jusczyk, Houston and Goodman (1998) for a review). Furthermore, it has long been observed that the perception of non-native stimuli strongly suggests that listeners have the expectation that the stimuli they hear are utterances in their L1. Such an L1 biased expectation seems to arise due to our internalized knowledge about the L1 sounds and their patterning, that is the phonology of our L1, which operates as a "filter" both in production and perception of non-native sounds.

One such kind of phonological knowledge pertains to phonotactics. Phonotactics broadly refers to our knowledge about which sound sequences are legitimate and which are not. Languages not only differ in the number and types of sounds they possess in their inventories but also in terms of the various possibilities that exist for the combination of the sounds they have. For instance, while German, English and Dutch allow consonantal sequences such as [dr] and [pl] to begin words, Turkish lacks such complex clusters in the same position although the language contains all of the relevant consonants in its phonemic inventory. The influence of L1 phonotactic restrictions is often reflected in the production of loan words, which are typically altered to fit the L1 norms. For instance, Japanese speakers produce English words that contain consonant clusters by epenthesizing into illicit consonantal sequences (e.g., [makɯdonarɯdo] 'Mac Donald', [sɯtoraikɯ] 'strike'). Japanese speakers' modification of L2 forms is perhaps not surprising given the very poor consonant cluster inventory of Japanese, which predominantly consists of CV syllables.

Research by Dupoux and colleagues (Dupoux et al., 1999) investigated an interesting question: are the epenthetic vowels added only in production, or are they in fact perceptually derived. They compared Japanese listeners with French listeners in their perception of consonant clusters in a series of behavioral experiments. An off-line phoneme detection task in their study employed nonce words such as [ebuzo] and [egudo] and they created series of six items from each by gradually reducing the duration of the vowel [u] down to zero milliseconds. The subjects were instructed to respond whether each item they heard contained the sound [u]. Unlike French listeners, Japanese listeners overwhelmingly judged that the vowel was present at all levels of vowel length. Interestingly, this was the case 70% of the time even when the vowel had been completely removed (i.e. the 0 ms condition).

Based on these and other similar results, Dupoux and colleagues concluded that the way in which a continuous speech stream is processed in the mind can be heavily influenced by the typical patterning of sounds in the L1 of the listener. Crucially, the influence can be so robust that listeners "invent" illusory vowels to accommodate illicit sequences of segments in their L1. Follow-up studies on Japanese employing a lexical decision task (Dupoux et al., 2001) and a neurolinguistic study using event-related potentials (Dehaene-Lambertz et al., 2000) have further confirmed the perceptual epenthesis phenomena in Japanese subjects.

There are several reasons, however, why a comparison between Japanese and French is not ideal. Japanese contains very few coda-onset clusters as it is predominantly a CV language, and only licenses coda consonants under a very restricted set of circumstances. Specifically, a consonant can only occur in the coda position if it forms the first member of a geminate construction (1a, b), or if it is a homorganic nasal (1c, d) (Itô, 1986; 1989). Coda consonants (other than the mora nasal) not carrying either of these two properties cannot occur in Japanese, as can be seen in the hypothetical examples given in (2).

(1)	a.	kap.pa	'a legendary being'	(2) a.	* kap.ta
	b.	gak.koo	'school'	b.	* tob.ba
	c.	tom.bo	'dragonfly'	c.	* pa.kat
	d.	kaŋ.gae	'thought'		

Closer examination on the stimulus items used in Dupoux et al. (1999) reveals that all the test words contained illicit coda consonants (shown in bold): e.g., [i**ʃ**.to], [e**b**.zo], [e**k**.to], etc. Therefore, we cannot tell whether the perceptual epenthesis induced in the percept of words such as [ebzo] is due to (1) contact restrictions (i.e., the sequence [b.z] is impossible), or (2) coda restrictions (i.e., [b] cannot be an independent coda). Since Dupoux and colleagues' primary aim in the study was to document the perceptual vowel epenthesis itself, they did not consider that two phonological explanations for their findings are confounded by the restrictedness of Japanese phonology. Consequently, they chose a more general explanation for Japanese speakers' perception of epenthetic vowels as being the result of their specific knowledge about which consonantal *sequences* are licit and which are not in the L1. That is, their explanation was in terms of contact.

In the present study, we aim to tease apart consonantal contact restrictions from coda restrictions, the confounded factors in the previous studies on Japanese, by employing Korean, which offers a much more interesting array of consonant clusters.

2. Korean Phonotactics

Korean exhibits a three-way contrast for oral stops and affricates: plain (C), aspirated (C^h) and tense (C'). Like Japanese, Korean does not allow complex coda and onsets. However, unlike Japanese, Korean offers a number of consonantal contact situations that help us to separate coda and contact restrictions. First, Korean puts an absolute ban on certain consonants in coda position (e.g., *[c.], *[s.], *[ɾ.], *[h])[1]. Second, it additionally disallows certain consonant clusters to surface. For instance, while the sequences *[k.m] and *[l.n] are illicit, [k.t] and [l.t] are licit. As in Japanese, Korean speakers produce foreign words that contain consonantal sequences deviating from Korean phonotactics with epenthetic vowels; [i] after palatal consonants and [ɯ] in other contexts (e.g., [a.i.sɯ.kɯ.ɾim] 'ice cream'; [sɯ.pʰən.ji] 'sponge').

The illicit codas and consonantal contacts are repaired by various phonological rules. In the following, we present only those processes that are relevant for the present study. First, strident consonants such as [c], [cʰ], and [s] neutralize in codas to the unreleased stop [t]. For instance, morphophonemic forms such as /nac/ 'daytime', /nacʰ/ 'face' and /nas/ 'sickle' become homophonous when they are pronounced in isolation (i.e., [nat]). Second, Korean has a nasalization rule that turns stops into nasals in pre-nasal contexts (e.g., /k.m/ → [ŋ.m]: /hak+mun/ → [haŋ.mun] 'learning'; /p.m/ → [m.m]: /cip+mun/ → [cim.mun] 'house gate'). Third, a process known as lateralization affects nasal sounds after lateral sounds (/l.n/ → [ll]: /tal+naɾa/ → [tal.la.ɾa] 'moon country'). Finally, voicing in Korean is predictable. Plain sounds become voiced intersonorantly (e.g., /pa+po/ → [pa.bo] 'idiot'). It should be noted that all of the processes illustrated above apply across words (though not across intonational phrases) without any restrictions in Korean.

3. Research Questions

In summary, there are at least three reasons why a word in the form of *[$VC_1.C_2V$] can be impermissible in Korean. First, [$C_1.C_2$] could be an illicit

[1] In this paper, we adhere to the phonetic symbols commonly used in the Korean linguistics literature. Specifically, [c] and [j] stand for voiceless and voiced alveo-palatal affricate, respectively. The symbol [ɯ] represents the high back unrounded vowel.

contact (e.g., *[k.m], *[l.n]), in the present cases triggering either nasalization or lateralization in the Korean production grammar, as explained above. Second, [C_1] could be an illicit coda (e.g., *[c.]). Third, [C_2] could be an illicit onset (e.g., *[.l], *[ŋ]). In this paper we will only be concerned with the first two factors, see Kabak (2003) for additional discussion of consonants disallowed in onsets.

We investigated whether coda or contact violations caused perceptual epenthesis. We also investigated whether contact violations (e.g., *[k.m], *[l.n]) were perceptually altered to fit their likely surface forms according to the norms of the Korean production grammar (that is, for example, whether words with [k.m] were confused with words with [ŋ.m]). Furthermore, we looked at the phonological status of the violation; namely whether perceptual epenthesis was induced only by contrastive features of Korean (e.g. [strident], [nasal]) or whether it could also be induced by allophonic features (e.g. [voice]).

3.1. Hypotheses

Given the ways in which *[C_1.C_2] sequences can be illicit in Korean, we propose two main hypotheses to be tested. First, the Consonantal Contact (CC) Hypothesis relates perceptual epenthesis to illicit sequences. Specifically, it asserts that Korean listeners hear epenthetic vowels when a given sequence is not valid in Korean. Second, the Coda/Onset Identity (COI) Hypothesis states that perceptual epenthesis arises when there is a syllable structure violation. Specifically, it predicts that Korean listeners hear epenthetic vowels when *[C_1.] is an illicit coda consonant. It should be noted that each hypothesis involves a different conception of perceptual epenthesis. While the CC hypothesis motivates perceptual epenthesis to break up illicit sequences of consonants, creating well-formed syllables is the primary goal of the COI hypothesis.

With regard to our secondary question as to the nature of L2 representations, we propose two sub-hypotheses. According to the Phonetic Processing Hypothesis, all phonetic features in the L1, including allophonic ones, such as voicing in Korean, are represented upon the perception of L2 forms. The Phonological Processing Hypothesis, on the other hand, claims that only contrastive (non-redundant) features in the L1 are represented in the L2 whereby predictable phonological information is suspended. Thus, the Phonetic Hypothesis claims that misplaced instances of [voiced] segments should also induce perceptual epenthesis.

3.2. Design

We employed an AX discrimination paradigm where pairs of nonce words with and without consonant clusters (e.g., [p^{h}a**k**m**a**] vs. [p^{h}a**k^{h}ʊm**a]) were compared with the assumption that if Korean listeners hear epenthetic vowels, they are likely to interpret the two words in the pair to be the same. Naturally produced words in the form of [p^{h}á**C$_1$(V)C$_2$**a] were used to construct test pairs. Following the Korean epenthesis patterns, the vowel was [i] after palatals and [ʊ], the closest approximation of [ɯ] elsewhere. C_1 varied between a permissible coda, [k] and [l], and an impermissible one, [c]. The onset of the second syllable, that is C_2, was either a stop ([t]) or a nasal ([m] or [n]). Consequently, the combination of C_1+C_2 produced either a permissible contact ([p^{h}á**kt^h**a], [p^{h}á**lt^h**a]) or an impermissible one (*[p^h**akma**], *[p^h**alna**]). It should be remembered that in the illicit contact cases, the coda consonants (i.e., [k] and [l]) are permissible coda consonants. The combinations of these consonants, however, with a following nasal can never surface in Korean due to (1) nasalization process in the case of [k], and (2) the lateralization process in the case of [l]. Rather, in the Korean production grammar, [k.m] and [l.n] surface as [ŋ.m] and [l.l], respectively. On the other hand, when C_1 is [c], neither a following stop (i.e., *[p^h**act^h**a]) nor a following nasal (*[p^h**acma**]) yield permissible sequences of consonants since [c] can never surface as a coda consonant in Korean.

We also employed the voiced counterparts of [k] and [c], that is [g] and [ɟ], respectively, in the same cluster combinations (i.e., [p^hagt^{h}a], [p^hagma]; [p^haɟt^{h}a], [p^haɟma]) to investigate specifically whether voicing, an allophonic feature predictable by the intervocalic voicing process in Korean, would be represented by Korean listeners, and whether voicing in an inappropriate context would cause perceptual epenthesis. Table 1 summarizes our test variables and their surface permissibility in the Korean production grammar.

Furthermore, the likely surface interpretations of *[p^hakma] and *[p^halna], that is [p^haŋma] and [p^halla]/[p^hanna], respectively, were also included in the experiment. This was specifically to test whether Korean listeners apply phonological rules such as nasalization and lateralization that are active in their production grammar to perceptually repair the illicit contacts *[k.m] and *[l.n].

Table 1: Status of test clusters in Korean

			C_2	
			Oral Stop (i.e., [t^h])	***Nasal (i.e., [n]or[m])***
C_1	**Licit**	**[k]**	Licit	Illicit
		[l]	Licit	Illicit
	Illicit	**[c]**	Illicit	Illicit
		[ɟ]	Illicit	Illicit
		[g]	Illicit	Illicit

AX test pairs were created by putting different exemplars of test words containing consonant clusters with words where the test clusters were separated by a vowel (e.g., [p^hacma] vs. [p^hac^{h}ima]; [p^hakt^{h}a] vs. [p^hak^{h}ʊta]), with a 1500 ms interval between each word . Based on the principles of the Signal Detection Theory (Green and Swets, 1974), we measured d', the ability to discriminate between same and different pairs. Thus, our test pairs also included different exemplars of the same words (e.g., [p^hak^{h}ʊma] vs. [p^hak^{h}ʊma]; [p^hakma] vs. [p^hakma]). An experimental block was created with 2 random repetitions of each of the test and filler pairs, yielding 118 trials in one block (39 test pairs plus 20 filler pairs presented in both possible orders). The blocks was presented 5 times each, yielding 10 total repetitions for each pair.

Twenty-five native speakers of Korean and 25 native speakers of English were recruited at the University of Delaware. All the Korean speakers were residing in the USA for educational purposes and none of them had started learning English before the age of 12. None of the English subjects knew Korean. Further details on the subjects are given in Kabak (2003).

3.3. Predictions

The predictions of the four combinations of hypotheses considered are summarized in Table 2, in which X indicates a situation where the subjects are predicted to be unable to discriminate the contrast between [paCCa] and [paCVCa].

Table 2: Discrimination predictions. X indicates a prediction of non-discrimination.

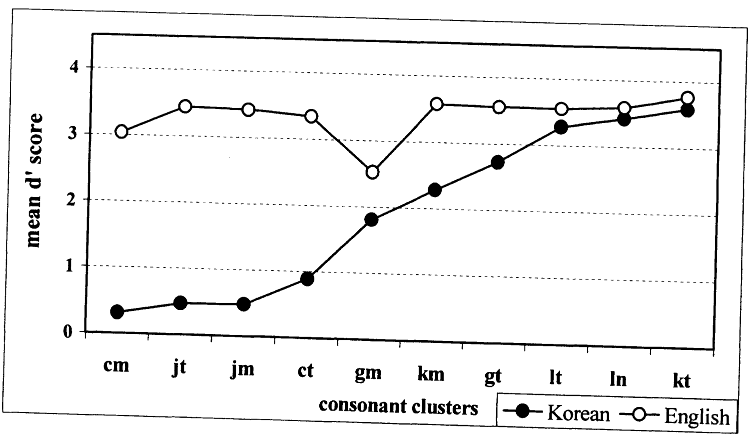
	CC Hypothesis		COI Hypothesis	
	Phonetic	**Phonological**	**Phonetic**	**Phonological**
k.t^h				
g.t^h	X		X	
k.m	X	X		
g.m	X	X	X	
l.n	X	X		
l.t^h				
c.t^h	X	X	X	X
c.m	X	X	X	X
j.t^h	X	X	X	X
j.m	X	X	X	X
	A	B	C	D

Under the Consonantal Contact (CC) hypothesis, all of the cases that yield illicit consonantal contacts, namely [k.m], [l.n], [c.m] and [c.t^h], should be misperceived by Koreans. If on the other hand, the Coda/Onset Identity (COI) hypothesis were true, all those with permissible coda consonants (i.e., [km], [kt^h], [ln], [lt^h]) should *not* be misperceived. When it comes to interpreting voiced consonants, if Korean listeners employ a phonetic processing approach and thus represent voicing then both [g] and [j] should behave like [c] since these consonants also never occur in coda position. On the other hand, under a phonological processing hypothesis, if voicing, being predictable, is perceptually suppressed, then [g] and [j] should behave like [k] and [c], respectively.

4. Results

The experimental conditions yielded d' values ranging from 4.21 (best) to 0 (worst). The average d' scores for both groups for the various clusters are shown in Figure 1. The English group successful discriminated all test clusters, with an average d' of 3.38. Observationally, the Korean listeners' d' scores exhibited a three-way grouping: (1) [cm], [jm], [jt^h], [ct^h], where the mean d' values were below 1; (2) [gm], [km], [gt^h], which formed an intermediate category with d' scores roughly between 2 and 3; and (3) [ln], [lt^h], [kt^h], where the d' scores were very close to the English group's mean d' scores. The first group with strident C_1 (which cannot occur in the Korean coda position), had an average d' of 0.53, indicating an inability to discriminate. The Korean group has a *median* d' value of 0 these clusters, whereas for the other clusters the median d' scores were similar to those for the English group. Interestingly, both groups' performance on [gm] was somewhat degraded. While there is no immediate explanation for why both groups suffered some degree of difficulty with [gm], this performance was not nearly as bad as the Korean group's performance on strident clusters.

Figure 1: Mean d' scores



A repeated measures Analysis of Variance (ANOVA) of d' revealed statistically significant effects throughout the model. All main factors were significant: language, $F(1, 48)=184.627$, $p<.0001$, first consonant (C_1), $F(4, 45)=96.131$, $p<.0001$ and second consonant (C_2), $F(1, 48=55.555$, $p<.0001$. All two-way interactions were also significant, though C_2 by language was marginal: C_1 by language, $F(4, 45)=57.871$, $p<.0001$, C_2 by language, $F(1, 48)=4.159$, $p<.047$, and C_1 by C_2, $F(4, 45)=18.492$, $p<.0001$. Ordinarily, these two-way interactions could be taken as the first step evidence for an effect of consonantal contact. However, the three-way interaction for C_1 by C_2 by language is also significant, $F(4, 45)=6.08$, $p<.001$, and therefore we need to look more carefully at the post-hoc tests.

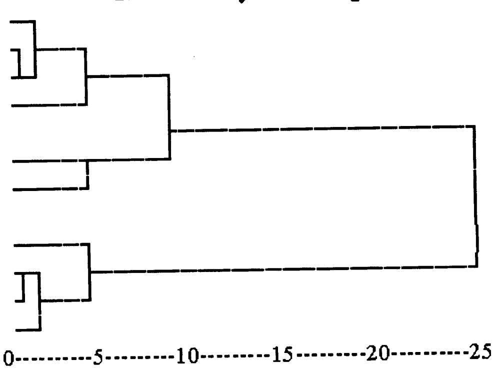
Scheffé and Tukey-Kramer post-hoc tests on the Korean group's d' scores consistently divided the consonant clusters into three groups (1) [cm], [ɟtʰ], [ɟm], [ctʰ], (2) [gm], [km], [gtʰ], (3) [ln], [ltʰ], [ktʰ], ordered from the lowest to the highest. With the data divided into three distinct groups, however, the main hypotheses of the study cannot be evaluated in a straightforward way. First, it cannot be said that the intermediate performance on [km], [gm] and [gtʰ] provides evidence for the CC hypothesis because the cluster [ln], which also induces a consonantal contact violation in Korean, is among the top three successfully discriminated clusters in the Korean group and overlaps with the English group's performance. Indeed, this finding, by itself, disputes the CC hypothesis. Second, the Korean group's performance on these clusters is not as bad as the performance on the strident clusters, which yielded near complete indiscriminability. Third, we constructed our hypotheses in such as way that they allow for only two distinct levels of performance on the test consonant clusters: (1) those that yield poor performance, and (2) those that yield good performance. Therefore, relative degrees of badness needed to be distinguished in order to better evaluate the hypotheses. The question, therefore, remained as to how the discriminability indices on the consonant clusters could be grouped if only two different groups of performance were to be made.

Using SPSS, we ran a multivariate hierarchical Cluster Analysis (CA) on the Korean data to answer this question. The CA produces a dendogram that graphically clusters cases (in this case, the test consonant clusters) by starting with single member clusters, which are then gradually fused until finally one large cluster is formed (see Figure 2). In a dendogram, the degree of association is strong between the members of the same cluster and weak between members of different clusters. The ruler at the bottom of the dendogram gives an index of an arbitrary measure of dissimilarity. Accordingly, from left to right, the CA divided the consonant clusters into 8, 6, 3, and 2 groups, then finally into one big group. While the 8 and 6-way clustering cannot be interpreted in any phonologically meaningful sense, the 3-way clustering, (1) [cm], [jt^{h}], [jm], [ct^{h}], (2) [gm], [km], (3) [gt^{h}], [ln], [lt^{h}], [kt^{h}], cannot sufficiently provide evidence against the COI hypothesis and the Phonological processing hypothesis because there is still a difference between the middle group (i.e., [gm] and [km]), and the lower group (the strident clusters). Furthermore, not all clusters with voiced segments are in the middle group (i.e., [gt^{h}] is grouped with the clusters that induced high d' scores), suggesting that the voicing information on C_1 does not matter for the Korean listeners, which rules out the Phonetic Processing hypothesis. As mentioned before, another cluster that induces a contact violation has been discriminated by the Korean listeners as well as the English group thus the middle group does not contain this cluster, either.

Figure 2: Korean d' cluster analysis

Sequence	**Type**	**mean d'**
kt	OK	3.58
ln	Bad contact	3.42
lt	OK	3.28
gt	Bad contact	2.72
km	Bad contact	2.28
gm	Bad contact	1.81
ct	Bad coda	0.89
jm	Bad coda	0.48
jt	Bad coda	0.47
cm	Bad coda	0.31

Cluster Analysis Dendogram



The two-way grouping that emerges out of the CA, on the other hand, groups the test clusters in a more homongeonus way, all the strident clusters on one side and the rest of the test clusters on the other. Such a grouping matches the predictions of Column D of the evaluation table provided in Table 2 above, where only strident codas are expected to be bad. This confirms only the Coda/Onset Identity (COI) hypothesis in combination with the Phonological processing hypothesis.

It should be noted that the [l]-clusters, namely, [ln] and [lt^h], form a replication of the same question that [km] and [kt^h] aimed to test in the study. We used these clusters as an independent test for the evaluation of the main hypotheses. According to the CC hypothesis, [ln] should be bad while both clusters should be successfully discriminated under the COI hypothesis since [l] is a legitimate coda consonant. An analysis of the lower and upper bounds for the clusters [ln] and [lt^h] within 95% confidence rate reveal that these clusters are not significantly different from each other. In addition, both groups do not differ from one another on these clusters. A power analysis reveals that any real difference in d' scores for these clusters must be less than 0.8 units of d' with 99% confidence if such a difference existed. With such small probable difference, the statistics on the [l]-clusters provide no support for the CC hypothesis, yet again supporting the COI hypothesis. Additional statistical analyses can be found in Kabak (2003).

The experiment also included a number of doublets which compared [km] and [ln] with their likely output forms in the Korean production grammar, [ŋm], [ll]/[nn], respectively. This was done to guard the conclusions on perceptual epenthesis from the effects of other types of phonological adjustments such as nasalization and lateralization on consonant clusters in Korean. The mean d' scores for these pairs revealed that they are all relatively similar and very high in both the English (mean d'= 3.72) and the Korean (mean d'= 3.59) groups, indicating that the Korean listeners did not confuse the illicit consonant clusters with their likely output forms in Korean. This suggested that not all phonological processes are relevant for the Korean listeners' perception, an important finding that provides evidence against the P-map hypothesis (Steriade, 2001a), which predicts perceptual similarity between underlying forms and output forms that are derived from them. Our study suggested that the difference between [km] and [ŋm], and [ln] and [ll] was very noticeable to Korean listeners; nevertheless, the alternation from these input forms to output forms is made in Korean production.

In summary, the results indicated that a given consonantal sequence C_1C_2 containing permissible coda consonants as C_1 is distinguishable from its epenthetically adjusted counterpart (i.e., C_1VC_2) for the Korean group regardless of whether C_2 is a plosive or a nasal consonant. As a consequence, some clusters (e.g. [km]) that do not occur in Korean can nevertheless be accurately perceived by Korean listeners. This supports the syllabically conditioned COI hypothesis in that the response patterns can only be explained if we assume that the L1 syllable structure violations, rather than contact violations, play an important role in the perception of consonant clusters. Furthermore, the voiced coda consonants do not cause perceptual epenthesis in the Korean group, suggesting that the voicing information is suppressed in speech perception. This supports phonological processing of features based on their abstract underspecified representations in the L1 of the listener.

5. Discussion

In this paper, we investigated two competing hypotheses based on two different views on phonotactics. The Consonantal Contact hypothesis followed a string-based approach towards phonotactics, predicting perceptual epenthesis as a consequence of consonantal contact restrictions in the L1 of the listener. The Coda/Onset Identity hypothesis assumed that perceptual epenthesis arises if the illicit consonantal sequence violates the L1 syllable structure conditions. Three different levels of performance on the test clusters were observed. Various statistical analyses suggest that the discrimination of the consonant clusters with strident consonants, namely [c] and [j], was significantly worse than other clusters with [g], [k], and [l] as C_1. Based on empirical evidence, we claim that having a sequentially illicit consonantal sequence is not sufficient to induce perceptual epenthesis. Rather, the cluster must induce a syllable structure violation in order to evoke perceptual epenthesis effects. Our results constitute substantial evidence against any analyses that employ syllable-independent, string-based and linear statements to explain consonantal phonotactics (e.g., Steriade, 1999; 2001a; Blevins, 2002).

The present study also revealed that voicing information was suppressed by the Korean group. Specifically, the clusters [gt^h] and [gm] were successfully discriminated while [jt^h] and [jm] were treated just like [c]-clusters. We suggest that [j] was mapped to /c/, therefore, it induced perceptual epenthesis because the segment is still a strident, whereas [g] was mapped onto /k/, which is a licit coda consonant. We take this finding to suggest that L2 representations are constructed on the basis of the abstract phonological properties of the L1 system of contrast and that the perceptual system suspends featural information in the speech signal if the detected values correspond to those that are underspecified in the L1 of the listener.

We obtained another interesting finding suggesting that not all phonological processes are relevant for speech perception. The experiment showed that processes such as nasalization and lateralization do not constitute possible perceptual strategies for Korean listeners. Specifically, although the successive occurrences of [k] plus [m], and [l] plus [n] are banned on the surface and /km/ → [ŋm] and /ln/ → [ll] are all mandatory in Korean production, the processes of assimilation do not play the same role as epenthesis in production. In the case of [strident] codas, which are similarly impossible on the surface, Korean listeners employed perceptual epenthesis to alter the input. The perceptual strategy that they employed is not in line with the way the same feature is accommodated in the production grammar. While in perception coda [strident] sounds were heard in the onset position causing perceptual epenthesis, in production Korean delinks the same feature from the coda position (i.e., /c/ → [t]). In short, the perceptual system does not appear to use repair strategies that are the simple inversions of the assimilatory processes that are active in the production grammar. This finding provides evidence against models that incorporate listeners' knowledge of perceptibility of sound contrasts to predict phonological alternations.

One such view is employed by Steriade (2001b) in her P-map hypothesis, which attempts to explain phonological alterations based on listeners' observations that certain contrasts are more discriminable than others, and that the same contrast is more salient in some positions than others (p. 236). Steriade claims that "the likelihood that a lexical representation R will be realized as R' is a function of the perceived similarity between R and R'" (p. 222). Indeed, the degree of similarity between any two features, by definition, should be inversely related to their discriminability. Thus, Steriade's statement is empirically testable using the present methodology. The discriminabilty indices we obtained are essentially the inverse of her "perceived similarity". Accordingly, the probability of saying [ll] given /ln/ must be proportional to a suitable function of the reciprocal of the discriminability score on [ll] vs. [ln]. The same equation can also be applied to [km] vs. [ŋm]. Since the discriminability is the inverse of similarity, as $d' \rightarrow 0$, p(similarity) $\rightarrow 1$. Our results show that the Korean group's d' scores on these pairs are very high (3.56 in the case of [ll] vs. [ln], and 3.73 in the case of [km] vs. [ŋm]). Such high scores result in probability scores that are very close to 0, which contradicts Steriade's statement regarding perceptual similarity to input. Specifically, if assimilatory processes were predicted by an index of perceived similarity between the assimilated variant and the underlying representation, Korean listeners should confuse clusters such as [ln] with its likely output form [ll] in Korean. Furthermore, we observed that strident codas are perceptually altered via epenthesis by Korean listeners, which suggests that they are indistinguishable from onset stridents. However, no such phonological alteration is attested in the synchronic phonology of Korean.

Our results also constitute important evidence against views that attribute perceptual preference for certain consonant clusters to the frequency with which those clusters occur in the language. If perceptual epenthesis were a means by which the perceptual system biases processing of clusters that have zero frequency, then all the illicit consonant clusters in the present study would be more susceptible to epenthesis. Our findings, however, show that Korean listeners' exhibit poor performance on only a certain set of consonant clusters (i.e., the strident cases) although all the illicit consonantal sequences have zero frequency of occurrence in Korean production. This suggests that a phonological influence of L1 phonotactic knowledge, rather than an effect of frequency, best explain the Korean groups' performance. Furthermore, this goes along with the findings of Moreton (2002), who argues for structural differences, rather than frequency differences, influencing English listeners' perceptual biases against certain illicit consonant clusters such as [dl] compared to [bw], both of which have zero frequency in English.

6. Conclusions

Our findings from Korean disentangle the phonological confounds in the previous experiments on Japanese, and strongly indicate that perceptual epenthesis is induced by the violation of syllable structure conditions of the L1, rather than string-based contact restrictions. Furthermore, our findings suggest that L2 forms are perceived based on the abstract underspecified properties of features in the L1 system of contrast. Neither views that rely on listeners' knowledge of perceptability of sound contrasts nor frequency-based accounts can explain Korean listeners' speech perception.

References

Blevins, J. (2002). The independent nature of phonotactic constraints: an alternative to syllable-based approaches. In C. Féry and R. van de Weijer (eds.). *The Syllable in Optimality Theory*, Cambridge University Press.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 6, 1568-1578.

Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes, 16*: 491-505.

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience*, *12*, 635-647.

Green, D. M., and Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*. Huntington, NY: Krieger.

Itô, J. (1986). *Syllable theory in prosodic phonology*. PhD Dissertation, University of Massachussetts, Amherst.

Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory, 7*, 217-259.

Jusczyk, P., Houston, D. & Goodman, M. (1998). Speech perception during the first year. In A. Slater (ed.), *Perceptual development: Visual, Auditory, and Speech Perception in Infancy*. Psychology Press.

Kabak, B. (2003). *The Perceptual Processing of Second Language Consonant Clusters*. Ph.D. dissertation. University of Delaware.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*, 55-71.

Steriade, D. (1999). Alternatives to syllabic analysis of consonantal phonotactics. In O. Fujimura, B. Joseph, & B. Palek (eds.), *Proceedings of the 2nd Linguistics and Phonetics Conference*, 1998 (pp. 205-245). Prague: Charles University, Press.

Steriade, D. (2001a). The Phonology of Perceptibility Effects: The P-map and its consequences for constraint organization. Ms. UCLA.

Steriade, D. (2001b). Directional asymmetries in place assimilation: a perceptual account. In E. Hume & K. Johnson (eds.), *The Role of Speech Perception in Phonology*. (pp. 219-250). New York: Academic Press.

Baris Kabak
University of Konstanz
Department of Linguistics
Fach D 174
78457 Konstanz, Germany

Baris.Kabak@uni-konstanz.de

William Idsardi
University of Delaware
Department of Linguistics
46 E. Delaware Ave.
Newark, DE 11, USA

idsardi@udel.edu

The Meanings of Consonants

BARBARA F. KELLY [1,3], WILLIAM R. LEBEN [2,3] and ROBERT H. COHEN [3]
[1]*University of California, Santa Barbara,* [2]*Stanford University,* [3]*LexiconBranding, Inc.*

1. Background: Sound Symbolism

Sound and meaning are supposed to be associated only arbitrarily in language, and usually they are. In terms of sound-meaning correspondence it makes no difference whether I eat cooked rice, *bhat*, or *nasi* – it's all the same thing. Still, linguists going at least as far back as Plato's *Cratylus* have come up with exceptions to the arbitrary pairing of sound and word meaning.

The study we report here shows that certain consonants convey meanings quite reliably, and consonants with similar meanings group into standard natural classes. In English there are dozens of cases where word-initial consonants seem to contribute a similar element to word meanings, including the following: *peg, pierce, pike, point, poke, probe, prong, prod; swab, swat, sway, sweep, swerve, swing, swirl, swish, switch, swoop; fire, flame, flare, flash, flicker* (Bloomfield 1933). This has also been attested cross-linguistically. For example, Jespersen (1933) compiles cases from a variety of languages, and Bolinger (1967) considers the morphemic status of these word-initial consonants and consonant clusters. More recently, 'Sound Symbolism,' edited by Hinton, Nichols, and Ohala (1994), examines symbolism cross-linguistically in its many different manifestations.

Saussure (1915) introduced the influential notion of *l'arbitraire du signe*, which identified the arbitrary association between sound and meaning as a basic feature of language. Following Saussure, linguists have espoused the view that the relation between words and meanings is entirely independent of the functional or physical relations between the two (Hockett 1963, Firth 1948). While it is impossible to ignore the essential validity of Saussurean arbitrariness of the relation between sound and meaning, it is worthwhile to look for cases in which the canon of arbitrariness is relaxed. Such cases are interesting because they tend not to be isolated or idiosyncratic. As Ohala (1994) points out, they are often rooted in attributes of the sounds. For example, Malkiel (1990, 1994) suggests that sound symbolism is so entrenched in some languages that regular sound changes fail to alter sound symbolic words, thus motivating new historical sound adjustments. Others have focused on more specific sounds and their sound symbolic properties. For example, vowels have been found to have certain sound symbolic properties. In a variety of languages front vowels have been associated with properties of being thin, bright, flat, hard while back vowels have been associated with properties of being large in size, heavy, low-pitched, dark (Argelander 1927, Fischer-Jorgensen 1978, Flournoy 1893, Jakobsen 1978, Newman 1933).

Ohala (1994) offers a possible explanation for some of the major cross-linguistic sound similarities. He surveyed past experimental and anecdotal work on sound symbolism and found strong cross-linguistic similarities. One reason for these patterns, Ohala proposes, is what he terms the *frequency code*. This associates sounds with high fundamental frequency with meanings like smallness and qualities such as politeness, deference, or submission. In contrast, words that have low pitch signals and sounds with low fundamental frequency are associated with properties of largeness and assertiveness.

Remarkably little work of an experimental nature has investigated sound symbolism. In one of the few early experiments that have been performed Sapir (1929) tested relations between different vowels and size. In this small experiment he presented subjects with the words *mil* and *mal* and asked them to determine "Which is a large table and which is a small table?" Around 80% of respondents designated *mil* as being the small table and *mal* as being the large. Although there is a possibility that there was semantic interference in this finding due to the association of *mil-* with the fraction 'one-thousandth,' several other experiments have corroborated this finding with respect to front and back vowels.

One of the most detailed studies to examine the sound symbolic properties of consonant sounds was carried out by Greenberg and Jenkins (1996). Greenberg and Jenkins found that among other things subjects judged stops as being relatively abrupt (as opposed to continuous), tight (as opposed to loose), rugged (as opposed to delicate), and inhibited (as opposed to free). Further, subjects judged stops and fricatives as relatively harsh (as opposed to mellow), rough (as opposed to smooth), active (as opposed to passive), sharp (as opposed to dull), difficult (as opposed to easy), and angular (as opposed to rounded).

In other research, LaPolla (1994) asked English speakers with no knowledge of Chinese to make judgments about meanings of spoken words in Cantonese. LaPolla found that English speakers showed a consistency in their use of sounds to guess at the meanings of words in a language they did not know. Words with high pitch were associated with smaller sizes and low pitch words were associated with larger sizes. This finding was corroborated by a separate study in which Cantonese speakers with no knowledge of English were asked to make judgments about English words and also about words in dialects they did not know. Even though some of the early experimental studies have their flaws, many of the sound symbolic effects they found have been corroborated in the non-experimental literature, including many of the survey articles in Hinton, Nichols and Ohala.

One area that has been largely overlooked in studies of sound symbolism is the study of emotions or intrinsic human qualities and speech sounds. Neuropsychology research has shown that some emotions are basic, unlearned states (Scherer and Ekman 1984) that can be conceptualized in terms of polar opposites (Plutchik 1980). For example, happiness has a polar opposite of sadness, surprise is opposite anticipation, and acceptance has a polar opposite of disgust. Gradations of intensity lead to different emotions, such as from anger to the less intense annoyance or the more intense rage. Conceivably, emotional states may occur along a continuum on which happiness and sadness are relative to each other.

Emotions are among the most basic of human qualities. To date, sound symbolism studies have examined sound-meaning correspondences relating to size, speed, color, sound, gender, and fundamental frequency. However, no documented sound symbolism research has systematically and empirically tested relations between human qualities and speech sounds. If natural class properties have sound correlates, what about another class relevant to speakers – human qualities?

The current study is an investigation of relations between human qualities and English consonant sounds. The aims of the study are:

i. To examine sound symbolism judgments of English speech with respect to a set of human qualities.
ii. To determine which English sounds best communicate the following meanings: *sad, alive, insecure, daring.*

Our overall goal in carrying out this empirical research is to discover whether there are sound-meaning correspondences relating to human qualities. Further, we aim to confirm the observations from a variety of languages and across a range of studies that there is a basis for claims regarding systematic relationships between sound and meaning.

2. Method

In an earlier related study, reported briefly in Cohen (1995), we tested 150 speakers of Dutch, Polish, Japanese, Italian and American English. For each language, respondents filled out a questionnaire in which they compared pairs of nonsense names for some given quality. A sample question was, "Which name sounds faster, *Tiron* or *Piron*?" Some of the findings established in this research include the discovery that fricatives are sound symbolically faster than stops and that voiceless consonants are smaller and more luxurious than voiced consonants.

In the current project we set out to examine sound symbolic judgments of English speakers with respect to word pair choices and their meanings. In an initial pilot study we asked respondents to judge word pairs according to six basic emotions: sad, happy, fearsome, angry, disgusting, and surprising. Of these, respondents only felt comfortable judging happy and sad. The remaining emotions were not considered appropriate for a study of word pair difference as respondents had trouble relating them to inanimate objects. This meant that a study of basic emotions and word pairs was not likely to yield meaningful results. To move beyond this problem we ran a further set of pilot tests where respondents made choices according to a set of three human qualities: *daring*, *alive*, and *insecure*, as well as the emotion *sad* which had tested well in the previous study.

Participants
Four hundred and eighteen (418) native speakers of American English participated in the study. Research was carried out through the world-wide-web and subjects completed the study at home using their own computers. The experiment was conceived and designed by researchers at Lexicon Branding, Inc and fielded on the web through the auspices of NFO Research and their national consumer panel. Subjects were controlled for variables of age (all were over the age of 18), sex (207 female and 211 male) and geographical location within the US.

Procedure
Participants were administered an online survey of forced choice alternatives presenting binary choices linking sound features with particular meanings. They were asked about a pair of nonsense names differing only by a single phonetic feature of the initial consonant. Respondents were asked to judge which of the two nonsense names rated higher on a given human quality. For example, for each target quality respondents were asked questions such as: "Which sounds more daring, *Paressa* or *Taressa*?"

Questionnaires consisted of an initial hypothetical product description and 68 questions. Two groups of respondents were given surveys testing the same sound pairs with different words to partial out possible choices on the basis of semantics. For example, in testing liquids, Group 1 was given the pair Renno/Lenno and Group 2 was given the pair Renza/Lenza. In order to counter a skewing effect on the basis of presentation ordering all emotions presented, sound pairs tested, and initial phoneme presentations were randomly ordered throughout the study.

The sound features included in this study are a subset of the features used by linguists to describe and differentiate sounds across the world's languages. While all of the consonant sounds examined here have been investigated in research surveyed above, in the current experiment they were systematically tested with respect to human qualities. The features and individual sound segments tested in the study were as follows:

VOICING:	voiced [b, d, g, v] vs. voiceless [p, t, k, f]
NASALITY:	nasal [m, n] vs. oral [b, d]
CONTINUANCY:	fricative [f, v, s, z] vs. stop [p, b, t, d]
PLACE:	labial [b, p] vs. dental [d, t] vs. velar [k, g]
LIQUIDS:	rhotic [r] vs. lateral [l]

In order to make the task more natural and meaningful, respondents were asked to interpret word pairs as names in the context of a performance sedan. All initial consonants were followed by vowels, and target words had either two or three syllables with CV(C) structure. On the basis of pilot tests the word *daring* was replaced by *courageous* with *d-* initial sounds in order to avoid interference due to alliteration. Other skewing due to frequency of sound and letter correspondences was controlled for by exchanging the letter *k* for c (pronounced [k]), since *k* occurs rarely in word-initial position in English. You may get a sense of the task faced by participants by imagining that these are names of cars and trying to answer questions like the following:

(1)

1.	Which sounds more alive?	Sekka	Zekka
2.	Which sounds more daring?	Balari	Palari
3.	Which sounds more sad?	Merrona	Berrona
4.	Which sounds more insecure?	Dobrax	Tobrax

3. Results

Results were tabulated by combining data for each of the questions dealing with a given sound contrast/attribute combination. This yielded results such as the overall percentage of respondents who selected target words with a voiced [d] versus a voiceless [t] for the question *Which sounds more insecure*?

People react differently to questions such as in (1) above. Some find the answers obvious, while for others they require extremely subtle judgments. Yet results were surprisingly consistent across the sample of 418 subjects. In order to determine if there were differences between the percentage of respondents choosing one sound segment over another across the different sets of word pairs, *t*-tests were carried out on all the data. Zero effects were found for age, sex, and location. Significant differences at the 90th (p=<.10) and 95th (p=<.05) percentile were found across all pairs excepting [d] vs. [n] in the category *insecure.*

Looking again at the emotions and words presented in (1) above we can see how the overall results patterned across some of the target words and emotions.

(2)

Which sounds more alive?	Sekka	**Zekka**
Which sounds more daring?	Balari	**Palari**
Which sounds more sad?	**Berrona**	Merrona
Which sounds more insecure?	**Dobrax**	Tobrax

The highlighted words in (2) are the words that were most often chosen as conveying the human quality targeted in the question. In the voiced pair [s, z] respondents selected [z] for the quality *alive*. For the quality *daring* the voiceless bilabial stop [p] was considered more daring than the voiced counterpart [b]. Subjects judged the oral bilabial stop [b] to be significantly more sad than the nasal bilabial [m], and respondents overwhelmingly selected the voiced stop [d] versus the voiceless stop [t] for the quality *insecure*.

Several clusters of features correlated with given symbolic meaning. For example, voicing and continuance combined to produce a very strong effect: [v] was found to be substantially more alive and daring than [b]. However, in non-continuant sounds the opposite effect was found whereby voiced stops [b, d, g] were less alive and daring than voiceless stops [p, t, k]. Overall, the features patterned consistently regardless of what other sound features they combined with. For example:

Voiceless stops sound more alive and daring than voiced stops.
Voiceless stops sound less sad and insecure than voiced stops.
Voiced fricatives sound more alive and daring than voiceless fricatives.
Voiced fricatives sound less sad and insecure than voiceless fricatives.
Dental consonants sound more alive and daring than non-dentals.

5. Discussion

The results fall into place so nicely that it is worth recalling our expectations before the experiment. We were prepared to discover that the consonants that best communicated the meaning *alive* were different from those that communicated *daring* and they had nothing to do with the results for *sad* and *insecure*.

Furthermore, we would not have been totally surprised if the set of consonants that best designated one human quality was arbitrary from a phonetic standpoint. For example, we may have found a pattern where [f, d, l] grouped together for one quality and [f, s, g] grouped together for another. Our study lends itself to several interpretations regarding what features are responsible for groupings. The patterning of voiced fricatives with voiceless stops and of voiceless fricatives with voiced stops is the most puzzling. Clearly, along with the standard features of voicing and manner of articulation, other phonetic factors are determining our results. We cannot speculate here on what those factors might be, but the groupings do have some importance, given the consistency of the results across linguistic features and across attributes of human qualities.

The identical patterning of *alive* and *daring* and of *sad* and *insecure* suggests that the attributes share a semantic relation. Conceivably, the pairs form some kind of semantic unit. This may be a unit as broadly defined as *positive* versus *negative*, with *alive* and *daring* being positive and *sad* and *insecure* being negative. It is fortunate that both *alive* and *daring* were included in the study as the performance of one constitutes an independent check on the other. The same applies to *sad* and *insecure*. They could have come out totally different from one another, but the fact that they didn't demonstrates a semantic relationship between the two attributes.

An even more unexpected finding is that the scale of consonants expressing *sad* and *insecure* is exactly the opposite of the one expressing *alive* and *daring*.

Some of the sound distinctions for the different human qualities are laid out below:

(3)

ALIVE / DARING				**SAD / INSECURE**
Voiceless stops	>	Voiced stops	>	Voiceless stops
Voiced fricatives	>	Voiceless fricatives	>	Voiced fricatives
Dental consonants	>	Non-dental consonants	>	Dental consonants
Stops	>	Fricatives	>	Stops

As we can see when we look at (3) above, voiceless stops were more *alive* and *daring* that voiced stops but voiced stops were more *sad* and *insecure* than voiceless stops. Again, we would not have predicted this patterning before the study. The exact relationship between the two scales provides yet another independent check on the validity of the study, since it is unlikely that one scale should turn out to be the reverse of the other purely by chance.

Since the study asked questions and elicited answers in written form, it is important to consider whether these results speak to sound at all. Conceivably, they might simply tell us about the meanings associated with written characters. If so, the results would be no less interesting, though it would then no longer be appropriate to interpret the experiment as a test of sound symbolism.

Subjects in this study read a written word, said the word aloud, and then made a judgment regarding the relation between the word and a given human quality. Since the written medium played such an important role, it is unlikely that only sound affected the choices that were made. More likely, respondents based their choices on some combination of phonetic and letter-shape information. What is quite clear is that it was not only letter shapes that drove the results.

While the chosen methodology leaves somewhat open the relative roles of sound and writing in subjects' judgments, there are other senses in which the methodology seems to have been well chosen. The research is more easily replicable, and the pool of subjects used was larger than would have been the case if the procedure had been conducted orally. The consistency in responses across the 418 subjects suggests that the methodology did manage to establish a valid association between consonants and human qualities.

This study of the relation between human qualities and English consonant sounds adds to the small body of experimental work on consonant sounds and meanings. The results point to a scale of consonants expressing more versus less *alive*, *daring*, *sad*, and *insecure*. This study provides some systematic, empirical evidence for a connection between English consonant sounds and human qualities. It seems then, that although Cratylus, with his voiceless velar-initial name, was being a little daring when he overstated the naturalness of connections between a word and the thing it names, some consonants do have meanings.

References

Argelander, A. 1927. *Das Farbenhören und der synästhetische Faktor der Wahrnehmung.* Jena: G. Fisher.

Bloomfield 1933. *Language*. New York, Henry Holt.

Bolinger, D. 1967. *Aspects of language*. New York. Harcourt, Brace, Jovanovich.

Cohen, R. 1995. There's more to a name. *Stanford Business School Magazine* March edition, pp 8-9.

Firth, J R 1948. Sounds and Prosodies. In F R Palmer, ed. 1970. *Prosodic Analysis*. Oxford: Oxford University Press, 1-26.

Fischer-Jorgensen, E. 1978. On the universal character of phonetic symbolism with special reference to vowels. *Studia Linguistica* 32.80-90.

Flournoy, T. 1893. *Les phénomènes de synopsie.* Paris-Genève.

Greenberg, J. H., and J. J. Jenkins. 1996. Studies in the psychological correlates of the sound system of American English: III. Descriptive ratings of selected consonants. *Word* 22.207-242.

Hinton, L. Nichols, J and Ohala, J. J. 1994. *Sound Symbolism*, Cambridge University Press.

Hockett, C. 1963. Problem of universals in Language. In J. Greenberg (ed) *Universals in Language,* Cambridge MA pp. 1-29

Izard, C.E. 1991. *The psychology of emotions*. New York: Plenum.

Jakobson, Roman. 1978. *Sound and Meaning*, Cambridge, MIT Press.

Jespersen, Otto. 1933. The Symbolic Value of the Vowel 'i'. *Linguistica*, College Park, MD. Reprinted in Selected Papers of O. Jespersen in English French and German,

LaPolla, R. J. 1994. An experimental investigation into phonetic symbolism as it relates to Mandarin Chinese. In Hinton, Nichols, and Ohala, pp.130-147.

Malkiel, Y. 1990. *Diachronic Problems in Phonosymbolism,* Philadelphia, J. Benjamins,

Malkiel, Y. 1994. Regular sound development, phonosymbolic orchestration, disambiguation of homonyms. In Hinton, Nichols, and Ohala, pp. 207-221.

Newman, S. S. 1933. Further experiments in phonetic symbolism. *American Jnl of Psychology 45.53-75.*

Ohala, J. J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. HNO, 325-347.

Ortony, A., & Turner, T. J. 1990). What's basic about basic emotions? *Psychological Review*, 97, 315-331.

Plutchik, R. 1980. A general psychoevolutionary theory of emotion. In R.
Sapir, E. 1929. A study in experimental symbolism. *Jnl. of Experimental Psychology* 12.225-239.

Saussure, Ferdinand de. 1966 [1915] *Course in General Linguistics.* New York: McGraw-Hill.

Scherer, K and P. Ekman, eds. 1984. *Approaches to emotion*, pp. 319-343. Hillsdale, N.J.: Lawrence Erlbaum.

Barbara F. Kelly
Dept. of Linguistics,
University of California,
Santa Barbara,
CA, 93106-3000

bfk0@umail.ucsb.edu

William R. Leben
Dept. Linguistics,
Stanford University
Stanford,
CA, 94305-2150

leben@stanford.edu

Morphologically Motivated Prosodic and Metrical Structures[*]

EUN-SOOK KIM
University of British Columbia

0. Introduction

In Nuu-chah-nulth some suffixes trigger alternation in stem vowels in length. There are three kinds of vowel alternation: vowel lengthening, vowel shortening and vowel alternation occurring with reduplication. In this paper, I provide a unified system to a superficially complex array of these apparently separate processes under Optimality Theory (McCarthy & Prince 1993, Prince & Smolensky 1993).[1]

1. Vowel lengthening

Some suffixes trigger lengthening of a stem-initial vowel. There are two patterns in this process: i) with some suffixes, if the second syllable of the stem is long, then it is shortened, and ii) with some other suffixes, only the first syllable of the stem is affected. In both cases, the second syllable does not have to be part of a root morpheme. In the case that the triggering suffix itself occupies the second syllable, the vowel of the suffix is not affected. (1) summarises the observation.

(1) Patterns of Nuu-chah-nulth vowel lengthening

Type	Triggering suffixes	Root/stem	
		1st syllable	2nd syllable
Type I	-(q)iiɬ, -ḥwaaɬ, -ʔiƛ, -, -ʔinḥi -p̓a(a), -p̓ičḥ	Lengthened	Shortened
Type II	-in̓ak̓uuḥ, -awiɬ, -panač,	Lengthened	No change

[*] I would like to thank my language consultants Mary Jane Dick, Sarah Webster, and Katie Fraser for sharing their language with me and for their enthusiasm and patience. I am also grateful to Marion Caldecott, Doug Pulleyblank, Pat Shaw, and Joe Stemberger for their insightful suggestions and corrections. This research is supported by the Jacobs Research Fund, and the Phillips Fund for Native American Research awarded to the author, and Hampton Research Grant awarded to Henry Davis.

[1] Nuu-chah-nulth, along with Ditidaht and Makah, constitutes the Southern Wakashan branch of the Wakashan language family. It is spoken along the west coast of Vancouver Island in B.C., Canada. The data in this paper are from Ahousaht, one of the 12 dialects.

Consider the following examples. *-(q)iiƚ* 'to make' is a Type I case. The first syllable of the stem is lengthened, while as seen in (2b), if the second syllable is underlyingly long, it is shortened on the surface.

(2) -(q)iiƚ 'to make'[2]

a. č'/a/pac-iiƚ	➔	č'[aa]paciiƚ
canoe-to make		'making a canoe'
b. s/a/pn/ii/-qiiƚ	➔	s[aa]pn[i]qiiƚ
bread-to make'		'making bread'

(3) is one of type II cases. The first syllable of the stem is lengthened; the second syllable, as seen in (3b), is not affected unlike type I.

(3) -panaj 'moving around'

a. ʔ/a/ya-panač	➔	ʔ[aa]yapanač
many-moving around		'Many people moving around'
b. n/a/qč/uu/-panač	➔	n[aa]qč[uu]panač
drunk-moving around		'(s.o.) moving around drunk (from place to place)'

2. Vowel shortening

Some suffixes cause the syllable(s) of a root or stem to shorten, which is the opposite of the lengthening process shown above.[3]

(4)

a. t/uu/ḥuk^w-(q)aq-mit-siš	➔	t[u]ḥuk^{w}aqitsiš
scared-very-PAST-1sg/IND		'I was very scared.'
b. č̓/ii/q/aa/-qaq	➔	č̓[i]q[a]qaqʔiš naʔaat
a spiritual song-very		'S/he is singing a spiritual song very loudly.

3. Vowel alternation with reduplication

Some suffixes trigger reduplication and sometimes the stem vowels are modified. There are 9 patterns in Nuu-chah-nulth reduplication in terms of the forms of both the reduplicant and the base. I provide 4 patterns in (5), which exhibit vowel alternation in the base (see Kim *in prep.* for the full discussion).

---

[2] The suffix initial /q/ disappears on the surface, when it follows a consonant-final stem.

[3] Abbreviations used in this paper are: DUR=durative, IND=indicative, INT=interrogative NEG=negative, RED=reduplicant, sg=singular, pl=plural.

(5) Patterns of Nuu-chah-nulth reduplication

TYPE	Vowel length in		RED-BASE
	Reduplicant	Base	
Class IV	Long	Affected; lengthened, if underlyingly short	CVV(C)-CVV(C)
Class V	Short	Affected; shortened, if underlyingly long	CV-CV(C)
Class VI	Long	Affected; shortened, if underlyingly long	CVV-CV(C)
Class VII	Short	Affected; lengthened, if underlyingly short	CV(C)-CVV(C)

3.1. Class IV: Red=σμμ; Base=$_1$σμμ

The reduplicant vowel is always long, and the base (or the first syllable of the base, if it consists of more than one syllable) is also long. If the vowel of the base is short, then it is lengthened as shown in (6).

(6) a. w̓aaw̓aasaqaʔiš
    RED-w̓asaq-(y)a-ʔiš
    RED-to cough-continuously-3sg/IND
    'She is continuously coughing.'

b. cuuccuucaʔiš
    RED-cuc-(y)a-ʔiš
    RED-to scratch-continuously-3sg/IND
    'S/he is continuously scratching.'

3.2. Class V: Red=σμ; Base=$_1$σμ

The reduplicant has a short vowel and (the first syllable of) the base is also short. If the vowel of the base is long, then it is shortened as shown in (7).

(7) a. ʔuʔusumk̓ukʔiš
    RED-ʔusum-k̓uk^{w}-ʔiš
    RED-to need/want-to resemble-3sg/IND
    'S/he appears to need (s.t.)'

b. ƛ̓iƛ̓ix^{w}ak̓uk
    RED-ƛ̓iix^{w}-(a)a-k̓uk^{w}
    RED-to smile/laugh-DUR-to resemble
    'Smirk'

3.3. Class VI: Red=σμμ; Base=$_1$σμ

The reduplicant is always long, but (the first syllable of) the base is short. If the vowel of the base is long, then it is shortened as shown in (8).

(8) a. wiiwikitẏak
RED-wik-itẏak
RED-NEG-afraid/fear
'Not afraid of anything'

b. siisicitẏaksiš
RED-siic-itẏak-siš
RED-maggot-afraid/fear
'I am afraid of maggots.'

3.4. Class VII: Red=σμ; Base=$_{1}$σμμ

The process is exactly opposite to Class VI reduplication: the reduplicant is always short, but if (the first syllable of) the base is underlyingly short, then it is lengthened as shown in (9).

(9) a. ʔuʔuušsapiʔiš
RED-ʔuuš-sapi-ʔiš
RED-some-to depend on-3sg/IND
'S/he is depending on someone.'

b. ʔaʔaaqisapiḥsuu	waaɬak	mituuni
RED-ʔaqi-sapi-ḥsuu	waɬ-ak	mituuni
RED-what-to depend-2pl/INT	to go-DUR	Victoria

'What are you depending on to go to Victoria?'

4. Analysis

We have seen so far that vowel alternation in Nuu-chah-nulth exhibits complex properties both phonologically and morphologically. I summarise the questions these data raise as follows:

1. How do we treat the different patterns of vowel alternations?
2. How do we treat the modification of base forms in terms of vowel length?
3. Is there any way of dealing with these processes under a unified system?
4. How do we integrate morphological aspects of the processes with phonology?

In the following sections, I will discuss these problems.

4.1. Metrical requirements

Modification of root/stem vowel length is due to metrical conditions specified for each vowel lengthening, shortening, and reduplication-triggering suffix. Adapting Pulleyblank *to appear,* I propose that metrical requirements are defined as templates as seen in (10) (cf. Crowhurst 1991, Hayes 1994).[4]

[4] See Kim (in prep.) for detailed discussion for Nuu-chah-nulth foot structure.

(10) I. Trochaic foot

a. $\text{FootForm}_{\text{I}}$: two light syllables $(\sigma \quad \sigma)_{\varphi}$

$\mu \quad \mu$

b. $\text{FootForm}_{\text{II}}$: two moras $(\ \sigma\ )_{\varphi}$

$\mu \quad \mu$

c. $\text{FootForm}_{\text{III}}$: two syllables with one heavy $(\sigma \quad \sigma)_{\varphi}$

$\mu \quad \mu \quad \mu$

d. $\text{FootForm}_{\text{IV}}$: two heavy syllables $(\sigma \quad \sigma)_{\varphi}$

$\mu \quad \mu\mu \quad \mu$

II. Iambic foot

e. $\text{FootForm}_{\text{V}}$: two syllables with first light: $(\sigma \quad \sigma)_{\varphi}$

μ

f. $\text{F(oot)F(orm)}_{\text{VI}}$: two moras: $(\ \sigma\ )_{\varphi}$

$\mu \quad \mu$

g. $\text{FootForm}_{\text{VII}}$: two syllables with first light; second heavy: $(\sigma \quad \sigma)_{\varphi}$

$\mu \quad \mu \quad \mu$

Each template is specified for specific suffixes and their surface effects are realized by the following constraints, (11), and their language-specific ranking, (12).

(11) Constraints:

a. FootForm_{δ}: A foot must obey the prosodic requirement specified on suffixes, if any. (δ indicates morphological domains: each class of suffixes.)

b. MAX_{μ}: Moras in the input must have a correspondent in the output.

c. DEP_{μ}: Moras in the output must have a correspondent in the input.

(12) Ranking:

$\text{FootForm}_{\delta} \gg \text{MAX}_{\mu}, \text{DEP}_{\mu}$

The implication of the ranking is that observing the metrical requirements is more crucial than maintaining the quantity of the input vowel.

The effect of the ranking is illustrated by the tableaux below. The first case is vowel lengthening. (13) is an example of Type I lengthening, where a triggering suffix is specified for FootForm III: the first syllable of the foot is long and the second is short.

(13) Type I: $FootForm_{III}$: $(\sigma \quad \sigma)_{\varphi}$

μ μ μ

(14) is the relevant example and (15) is its tableau.

(14) s/a/pn/ii/-qiiɬ ➔ s[aa]pn[i]qiiɬ
bread-to make 'making bread'

(15) Tableau for (14)

sapnii-qiiɬ(FFIII)	$FootForm_{III}$	MAXIOμ	DEPIOμ
☞ a. $(saapni)_{\varphi}$qiiɬ		*	*
b. $(sapnii)_{\varphi}$qiiɬ	*!		
c. $(sapni)_{\varphi}$qiiɬ	*!	*	

(16) is a case of Type II lengthening, where the fist syllable of the stem is long, but the second is not affected.

(16) Type II: $FootForm_{II}$: $(\ \sigma\ )\varphi$

μ μ

(17) n/a/qj/uu/-panaj ➔ n[aa]qj[uu]panaj
drunk-moving around '(s.o.) moving around drunk '

(18) Tableau for (17)

naqčuu-panač (FFII)	$FootForm_{II}$	MAXIOμ	DEPIOμ
☞ a. $(naaq)_{\varphi}$čuupanač		*	
b. $(naaqčuu)_{\varphi}$panač	*!	*	
c. $(naqču)_{\varphi}$panač	*!		
d. $(naqčuu)_{\varphi}$panač	*!		

(19) is a case of vowel shortening, where the first two syllables of the stem are shortened. (20) and (21) are the relevant example and tableau, respectively.

(19) FootForm$_I$: $\begin{array}{cc} (\sigma & \sigma)_{\varphi} \\ | & | \\ \mu & \mu \end{array}$

(20) t/uu/ḥuk^{w}-(q)aq-mit-siš ➔ t[u]ḥuk^{w}aqitsiš
scared-very-PAST-1sg/IND 'I was very scared.'

(21) Tableau for (20)

tuuḥuk^{w}-aq(FFI)	FootForm$_I$	MAXIOμ	DEPIOμ
☞ a. (tuḥu)$_{\varphi}$k^{w}aq		*	
b. (tuuḥu)$_{\varphi}$k^{w}aq	*!	*	
c. (tuuḥuu)$_{\varphi}$k^{w}aq	*!		*
d. (tuḥuu)$_{\varphi}$k^{w}aq	*!	*	*

4.2. Prosodic characterization of reduplicants

In the same spirit, each reduplication-triggering suffix is not only specified for a metrical requirement on the foot structure, but also manifests a prosodic requirement σ as a cooccurring prefix, adapting Marantz 1982, McCarthy & Prince 1986, Downing 2000, 2001, Pulleyblank *to appear,* and following Kim *to appear*. For example, a suffix of class IV, *-ya*, could be represented in the lexicon as follows.

(22) σ...ya

4.3. Indexation of faithfulness constraints

Adopting Itô & Mester (1999), I propose that phonological as well as morphological domain-specified faithfulness constraints are closely related to the modification of the base forms of some classes. That is, the following faithfulness constraints are domain-specified both phonologically and morphologically.

(23) Input-Output faithfulness: δ=phonological or morphological

a. MAXIO$_{\delta}$: Every segment of the input in the domain of δ has a correspondent in the output.

b. DEPIO$_{\delta}$: Every element in the output in the domain of δ has a correspondent in the input.

Segments are never deleted or inserted, so MAXIO/DEPIO(seg) are undominated. However, length alternation is suffix-dependant, leading to the following ranking:

(24) a. Classes IV & VII: FootForm >> DEPIO(μ)$_{IV,VII}$
b. Classes V & VI: FootForm >> MAXIO(μ)$_{V,VI}$

We need FaithBR constraints to obtain the effect of the Base-Reduplicant identity.

(25) Base-Reduplicant Faithfulness: δ=all classes (McCarthy & Prince 1994, 1995)
a. MAXBR$_{\delta}$: Every element of the base in the domain δ of has a correspondent in the reduplicant.
b. DEPBR$_{\delta}$: Every element of the reduplicant in the domain of δ has a correspondent in the base.

The coda of the base is copied depending on the triggering suffixes and the following markedness constraint plays a crucial role in determining the form of the reduplicant.

(26) Markedness
NOCODA: Syllables may not have codas.

(27) is the portion of the Nuu-chah-nulth grammar relevant for the particular processes under discussion:

(27) MAX/DEPIO(SEG), **MAX/DEPIO(μ)$_{I\text{-}III}$ MAXIO(μ)$_{IV,VII}$, DEPIO(μ)$_{V\text{-}VI}$**
⇩
FootForm$_{\delta}$
⇩
MAXIO(μ)$_{V,VI}$, DEPIO(μ)$_{IV,VII}$
⇩
MAX/DEPBR$_{[I\text{-}1,IV,\ VII\text{-}1]}$
⇩
NOCODA
⇩
MAX/DEPBR$_{[I\text{-}2,II,III,V,\ VII,\ VII\text{-}2]}$

The implication of the ranking is illustrated in tableaux as follows, with the relevant examples.

(28) Class IV suffixes- FootForm$_{IV}$: $(\sigma \quad \sigma)_{\varphi}$
μ μ μ μ

(29) cuuccuuca?iš
RED-cuc-(y)a-?iš
RED-to scratch-continuously-3sg/IND
'S/he is continuously scratching.'

(30) Tableau for (29)

/Rσ-cuc-(y)a(FFIV)/	**MAXIO IV**	**Foot Form$_{IV}$**	**DEPIOμ IV**	MAXBR DEPBR IV IV	NO CODA
☞a. {σ[cuuc]<u>cuu</u>}$_{\varphi}$<u>ca</u>			*		*
b. {σ[cuc]<u>cuu</u>}$_{\varphi}$ca		*!	*	*(μ)	*
c. {σ[cuc]<u>cu</u>}$_{\varphi}$ca		*!			*
d. {σ[cuuc]<u>cu</u>}$_{\varphi}$<u>ca</u>		*!		*(μ)	*
e. {σ[cuu]<u>cuu</u>}$_{\varphi}$<u>ca</u>			*	*!(seg)	
f. {σ[cuu]<u>cuu</u>}$_{\varphi}$ya	*!(seg)		*		

With FootForm$_{IV}$ outranking DEPIOμ, the base vowel is lengthened on the surface (and the reduplicant has a long vowel).

(31) Class V suffixes-FFI: $\begin{array}{cc}(\sigma & \sigma)_{\varphi} \\ | & | \\ \mu & \mu\end{array}$

(32) ƛ̓iƛ̓ixʷak̓uk
RED-ƛ̓iixʷ-(a)a-k̓uk
RED-to smile/laugh-DUR-to resemble
'Smirk'

(33) Tableau

/Rσ-ƛ̓iixʷ... k̓uk(FFI)	**MAXIO [Seg.]$_V$**	**Foot Form I**	**MAXIO μ$_V$**	NO CODA	MAXBR DEPBR V V
☞ a.{σ[ƛ̓i]<u>ƛ̓i</u>}$_{\varphi}$<u>xʷ</u>a			*		*(seg)
b. {σ[ƛ̓ii]<u>ƛ̓ii</u>}$_{\varphi}$<u>xʷ</u>a		*!			*(seg)
c. {σ[ƛ̓i]<u>ƛ̓ii</u>}$_{\varphi}$<u>xʷ</u>a		*!			*(seg) *(μ)
d.{σ[ƛ̓ixʷ]<u>ƛ̓i</u>}$_{\varphi}$<u>xʷ</u>a			*	*!	
e. {σ[ƛ̓i]<u>ƛ̓i</u>}$_{\varphi}$a	*!		*		
f. {σ[ƛ̓ii]<u>ƛ̓i</u>}$_{\varphi}$<u>xʷ</u>a		*!	*		*(seg) *(μ)

With $\text{FootForm}_{\text{I}}$ outranking MAXIOμ, the base vowel is shortened on the surface (and the reduplicant has a short vowel).

(34) Class VI suffixes-FFIII: (σ σ)$_{\varphi}$

μ μ μ

(35) siisicity̓aksiš
RED-siic-ity̓ak-siš
RED-maggot-afraid/fear
'I am afraid of maggots.'

(36) Tableau for (35)

/Rσ-siic-ity̓ak(FFIII)	**MAXIO [Seg.]$_{\text{VI}}$**	**Foot Form III**	**MAXIO μ$_{\text{VI}}$**	NO CODA	MAXBR VI	DEPBR VI
☞a. {σ[sii]si}$_{\varphi}$ci			*		*(seg)	*(μ)
b. {σ[sii]sii}$_{\varphi}$ci		*!			*	
c. {σ[si]si}$_{\varphi}$ci		*!	*		*	
d. {σ[siic]sii}$_{\varphi}$ci		*!		*		
e. {σ[sii]si}$_{\varphi}$ci	*!		*			*(μ)
f. {σ[sii]sii}$_{\varphi}$ci	*!	*				
g. {σ[si]sii}$_{\varphi}$ci		*!			*(seg) *(μ)	
h{σ[siic]si}$_{\varphi}$ci			*	*!		*(μ)

With $\text{FootForm}_{\text{III}}$ outranking MAXIOμ, the base vowel is shortened on the surface (and the reduplicant has a long vowel).

(37) Class VII suffixes-FFVII: (σ σ)φ

μ μ μ

(38) ḥicḥiicn̓uk
RED-ḥic-n̓uk
RED-feces-on the hand
'Feces on the hand'

(39) Tableau for (38)

/σ-ḥic-n̓uk(FFVII)	**MAXIO VII**	**Foot Form VII**	**DEPIO** μ_{VII}	MAXBR VII	DEPBR VII	NO CODA
☞ a. {σ[ḥic]$\underline{\text{hiic}}$}$_{\varphi}$			*	*(μ)		**
b. {σ[ḥii]$\underline{\text{hiic}}$}$_{\varphi}$		*!	*	*(seg)		*
c. {σ[ḥi]$\underline{\text{hic}}$}$_{\varphi}$		*!		*(seg)		*
d. {σ[ḥic]$\underline{\text{hic}}$}$_{\varphi}$		*!				**
e. {σ[ḥii]$\underline{\text{hi}}$}$_{\varphi}$	*!(seg)	*			*(μ)	
f. {σ[ḥii]$\underline{\text{hii}}$}$_{\varphi}$	*!(seg)	*	*			
g. {σ[ḥi]$\underline{\text{hiic}}$}$_{\varphi}$			*	*(seg) *!(μ)		*
h. {σ[ḥii]$\underline{\text{hic}}$}$_{\varphi}$		*!		*(seg)	*(μ)	*

With FootForm$_{VII}$ outranking DEPIOμ, the base vowel is lengthened on the surface (and the reduplicant has a short vowel).

5. Conclusion

Nuu-chah-nulth prosodic and metrical structures are morphologically determined. Vowel alternations motivated by vowel lengthening, shortening, and reduplication-triggering suffixes are due to metrical requirements specified for the suffixes. Moreover, multiple patterns of reduplication and modification of the base forms are due to such metrical requirements. Nuu-chah-nulth reduplication results from prosodic requirements specified for suffixes. In addition, Nuu-chah-nulth vowel alternations provide a case where faithfulness constraints can be both phonologically and morphologically domain-specified. This study provides a unified approach to a superficially complex array of apparently separate processes.

References

Crowhurst, Megan. 1991. *Minimality and Foot Structure in Metrical Phonology and Prosodic Morphology*. Ph.D. dissertation. University of Arizona, Tucson.

Downing, Laura. 2000. Morphological and prosodic constraints on Kinande verbal reduplication. *Phonology* 17: 1-38.

Downing, Laura. 2001. Ungeneralizable minimality in Ndebele. *Studies in African Linguistics* 30: 33-58.

Hayes, Bruce. 1994. Weight of CVC can be determined by context. In J. Cole & C. Kisseberth. (eds.).*Perspectives in Phonology*. 61-79. CSLI Publications.

Itô, Junko & Armin Mester. 1999. The structure of the phonological lexicon. In Tsujimura, Natsuko, ed. *The Handbook of Japanese linguistics*, Blackwell Publishers, Malden, MA, and Oxford, U.K. 61-100.

Kim, Eun-Sook. *To appear*. The patterns of reduplication in Nuu-chah-nulth. In M. Kadowaki & S. Kawahara, eds. *Proceedings of the NELS 33*. GLSA: University of Massachusetts, Amherst.

Kim, Eun-Sook. *In preparation. Theoretical issues in Nuu-chah-nulth phonology and morphology*. Ph.D. dissertation. University of British Columbia.

Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13: 483-545.

McCarthy, John & Alan Prince. 1986. Prosodic morphology. Ms. University of Massachusetts, Amherst & Brandeis University.

McCarthy, John & Alan Prince. 1993. Prosodic morphology I: constraint interaction and satisfaction. MS. University of Massachusetts, Amherst, and Rutgers University.

McCarthy, John & Alan Prince. 1994. The emergence of the unmarked: optimality in prosodic morphology. In *Proceedings of the NELS 24*, 333-379. GLSA: University of Massachusetts, Amherst.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk, eds. *Papers in Optimality Theory, University of Massachusetts Occasional Papers* 18: 249-384.

Prince, Alan & Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar. MS. Rutgers University and University of Colorado, Boulder.

Pulleyblank, Douglas. *To appear*. The patterns of reduplication in Yoruba. In Inkelas, S. & K. Hanson, eds. *The nature of the word: essays in honor of Paul Kiparsky*. The MIT Press.

Eun-Sook Kim
Buchanan E-270, 1866 Main Mall
Department of Linguistics
University of British Columbia
Vancouver, B.C.
Canada V6T 1Z1

eunsookk@interchange.ubc.ca

English Locative Inversion: Grammatical Interfaces and Constructions*

JONG-BOK KIM
Kyung Hee University

1. Introduction

Locative inversion (LI) is a general term applying to cases of inversion with the locative PP in the preverbal and the theme NP in the postverbal position. The examples of LI given in (1) are cases found in the ICE-GB (British Component of the International Corpus of English) that contains about one million words from 500 different spoken and written texts:

(1) a. And then suddenly from the bottom [appears] a motor car. (ICE-GB:S1B-038)
b. Then behind him [came] Eton Lad who fluttered. (ICE-GB:S2A-006)
c. Within this quotation are [engraved] astrological zodiacal symbols, clockwise from Pisces uppermost.(ICE-GB:W2A-040)
d. In the top drawer of her desk [lay] her letter of resignation from Jupiter Services.(ICE-GB:W2F-008)

As noted by Birner and Ward (1998), such examples have syntactically more basic counterparts differing not in truth conditions but only in the way the informational content is presented. Several attempts have been made to account for how the grammar makes it possible to say the same thing in different ways in such LI constructions, focusing on the following questions:[1]

*An earlier version of this paper was presented at the 29th Berkeley Linguistic Society at University of California, Berkeley at February 14–17, 2003. I thank Emily Bender, Chan Chung, Ivan Sag, Peter Sells, and the audiences of the 27th BLS for valuable comments and criticisms. Part of this paper has been developed from Chung and Kim (2002). All errors and misinterpretations are, of course, mine. I also wish to acknowledge the financial support of the Korea Research Foundation (KRF-2003-A00443).

[1] There have two main perspectives in the treatment of English LI: transformational and lexicalist. The transformational perspective has treated LI in terms of movement and configurational structures.

- How can we generate such locative constructions and what triggers the inversion?
- What are the grammatical properties of the preverbal PP and the postverbal NP, and the status of the verb in LI?
- How can we account for various grammatical (syntactic, semantic, and pragmatic) constraints in LI?

This paper is another attempt to answer these questions within the framework of constraint-based grammar, HPSG (Head-driven Phrase Structure Grammar). In particular, in order to account for intriguing and mixed properties of the constructions in question, the paper argues for a grammar that allows tight interactions among different grammatical components: valence, argument-structure, discourse-function (information structure), and constructions.

2. Mixed Properties of the Preverbal PP and the Postverbal NP

One of the controversial issues in English locative inversion (LI) is the functional status of the preverbal PP and the postverbal NP. As noted by Bresnan (1994) and many others, the preverbal PP functions both as a subject and topic, while the postverbal NP has subject as well as non-subject properties at the same time.

2.1. Subject and Topic Properties PP

Subject Properties: There seem to exist several phenomena that support the assumption that the preverbal PP is subject. In particular, as noted by Coopmans (1989), Bresnan (1994), Culicover and Levine (2001), among others, LI behaves just like subject with respect to phenomena such as raising, tag question, *that*-trace effect, weak cross-over effect, and *do* support phenomena. Let us compare canonical and LI sentences:

(2) a. An entire army of ants seems to have crawled over my windowsill.
b. There are whales in the ocean, aren't there?
c. That bunch of gorillas, Terry claims *(that) _ walked into the room.

This view can be at large divided into two main approaches: topicalization approach (cf. Bowers 1976, Rochemont and Culicover 1990), and unaccusative approach (cf. Coopmans 1989, Hoekstra and Mulder 1990). The basic idea of the topicalization approach is to move the PP into a topic position and the subject NP to a VP-adjoined position. Meanwhile, the unaccusative analysis takes the NP to be the object of an unaccusative verb and moves to the locative PP into a subject position. Meanwhile, the lexicalist perspective has been proposed in Levine (1989, 2002), Bresnan (1994), Green (1992), Kathol and Levine (1992) among others. The central assumption of this lexicalist perspective is that the lexical properties of the inversion verb play an important role in projecting LI sentences.

d. ??Who$_i$ does his$_i$ mother think [t$_i$ is genius]?
e. Which portrait of the artist hung/*did hang on the wall?

(3) a. Over my windowsill seems to have crawled an entire army of ants.
b. In the ocean are whales, aren't there?
c. Into the room Terry claims *(that) _ walked a bunch of gorillas.
d. Into every dog$_i$'s cage peered its$_i$ owner.
e. In which garden *did stand/stood a fountain?

The postverbal PP in (3)a functions as the raised subject just like the canonical NP subject in (2)a. The postverbal PP in (3)b also, like the subject NP in (2)b, functions as the target for the pronoun agreement in tag questions. Also, when the PP in LI is extracted from an embedded clause as in (3)c, the presence of an overt complementizer is obligatory: this is the situation we fi nd with the canonical NP subject in (2)c. In addition, though the movement of a *wh-* phrase to an A′-position induces weak cross-over effect as in (2)d, the postverbal PP exhibits no such an effect, implying that the PP is in a A-position. Finally, just as the *wh*-NP subject in (2)e, the *wh*-PP in (3)e does not allow the presence of the so-called dummy *do*.

Topic Properties: As argued in Schachter (1992), one strong argument for the topichood of the PP concerns the usage of an indefi nite PP. As shown in the contrast (4), it is possible to have an indefi nite locative proform in a noninverted sentence but not in an LI sentence:

(4) a. A child was found somewhere.
b. *Somewhere was found a child.

A topic element in general represents established information, and it is thus preferred to be defi nite. The unacceptability of the indefi nite locative phrase *somewhere* in (4)b could thus be attributed to the topichood of the PP in LI.

The PP also exhibits distributional similarities with a topic phrase. For example, just like a true topic phrase, the preverbal PP in LI ppears neither in a nonfi nite clause nor in an embedded *if* or *whether* clause as illustrated in (5):

(5) a. *Bill asked if [such books] John only reads at home.
b. *Bill asked if [near John's house] lies buried treasure.

The co-occurrence with the complementizer also indicates the topichood of the PP. When LI or topicalization occurs in an embedded clause, the overt complementizer *that* must appear, as shown in (6) (cf. Schachter 1992, Bresnan 1994, Ginzburg and Sag 2001):

(6) a. Mary said [*(that) the dog, the man kicked].
b. Mary said [*(that) under the tree sat a woman].

2.2. Subject, Object, and Focus Properties of the Postverbal NP

Subject Properties: The examples in (7) show that even though the postverbal NP is not in a canonical subject position, it plays the role of subject with respect to subject-verb agreement.

(7) a. In the garden stand/*stands two fountains.
b. Down through the hills and into the forest flows/*flow the little brook. (Levine 1989)

Object Properties: In LI, there is a strong tendency for an adverb to occur after the adjoined NP, implying that the NP is a direct object of the locative verb (Coopmans 1989, Penhallurick 1984, Kathol and Levine 1992):

(8) a. Into the room strode (*boldly) Robin boldly.
b. Outside the door sat (*uncomfortably) a young man uncomfortably.

Focus Properties: As noted by Rochemont (1986) the postverbal NP must be identifi ed as a presentational focus. Strong support can be found from the fact that LI sentences can freely alternate with their noinverted counterparts as responses to *wh* questions:

(9) A: Who ran into the forest?
B: Into the forest ran ROBIN HOOD.
B: ROBIN HOOD ran into the forest.

A similar case has been provided by Bresnan (1994). The focus property of the postverbal NP can explain why the reply B in (10) is in infelicitous.[2]

(10) A: I am looking for my friend Rose.
B: # Among the guests of honor was sitting Rose.

Since Rose is already mentioned in A, it cannot be reintroduced on the discourse as a presentational focus.

[2]Corpus research results also show that the postverbal NP represents new information, that is, information which the speaker assumes not be in the consciousness of the hearer at the time of the utterance. See Penhallurick (1984) and Birner and Ward (1998).

3. A Constraint-Based Approach

3.1. Constraints on the Mapping Relation

In generating English LI sentences, as a starting point we assume that that LI is a result of different mapping relations among grammatical function structure (valence structure), argument structure, and discourse structure. In particular, the paper takes discourse functions to be the main trigger for LI (cf. Schachter 1992 and Bresnan 1989, 1994). In canonical sentences with locative verbs, the theme NP is mapped onto SUBJ and the locative PP onto COMPS. However, as represented in (11), when the locative PP carries topic and the theme NP is assigned to have focus, this mapping relation is reversed:

(11) Mapping in the LI case:

Valence Structure:	SUBJ	COMPS
Argument-structure:	NP[theme]	PP[loc]
Discourse-structure:	Top	Focus

For example, a *loc-v* word like *sat* will take an NP and PP as its arguments and these are realized on the appropriate valence list (grammatical functions): SUBJ and COMPS as in (12)a.[3] However, as a *loc-inv* word, this mapping relation is reversed as in (12)b:

$$\text{(12) a.}\ \begin{bmatrix} \textit{loc-v-w} & \\ \text{PHON} & \langle \text{sit} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\ \text{NP}, \boxed{2}\ \text{PP} \rangle \end{bmatrix} \Rightarrow \text{b.}\ \begin{bmatrix} \textit{loc-inv-v} & \\ \text{PHON} & \langle \text{sit} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{2} \rangle \\ \text{COMPS} & \langle \boxed{1} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\ \text{NP}, \boxed{2}\ \text{PP} \rangle \end{bmatrix}$$

The reverse condition is possible only when the PP is mapped onto TOPIC and the NP is onto FOCUS, respectively. As a generalization, we could formulate this as

[3] It is hard to claim that LI is limited to unaccusative verbs. LI seems to include certain unergatives but exclude certain unaccusatives as shown in (i) (see Birner 1994, Levin and Rappaport 1995, Birner and Ward 1998):

(i) a. *On the top fbor of the skyscraper broke many windows.
  b. Behind the mayor's car marched police officers.
  b. On the third fbor worked two young women.

the following Locative Inversion Argument Realization Constraint:[4]

(13) Locative Inversion Argument Realization Constraint:

$$\textit{loc-v-lxm} \Rightarrow \begin{bmatrix} \textit{loc-inv-v} \\ \text{VALENCE} \begin{bmatrix} \text{SUBJ} \ \langle \boxed{2} \rangle \\ \text{COMPS} \ \langle \boxed{1} \rangle \end{bmatrix} \\ \text{ARG-ST} \ \langle \boxed{1}\ \text{NP}[\textit{can-ss}], \boxed{2}\ \text{PP} \rangle \\ \text{INFO-STR} \begin{bmatrix} \text{TOPIC} \ \langle \boxed{2} \rangle \\ \text{FOCUS} \ \langle \boxed{1} \rangle \end{bmatrix} \end{bmatrix}$$

Such dissociations between valence structure (grammatical functions) and argument structure motivated by the discourse functions directly reflect the basic properties of LI.[5] This analysis clearly indicates that the NP functions as subject at argument-structure level and an object at valence level, whereas the PP is a complement at argument-structure but a subject at valence-structure. Let us consider some immediate consequences of this line of approach.

Verb Types and Adjuncts The analysis fi rst does not allow transitive verbs to occur in LI as in (14):[6]

(14) a. *In the room rolled John the ball.
     b. *Onto the track ran the jockey the horse. (Rochemont 1986)

Transitive verbs like *roll* in (14)a are not *loc-v-lxm* and thus cannot serve as the input to the *loc-inv-v*. This also means that LI is possible only when the locative PP is a complement as noted in Bresnan (1994):

(15) a. ??/*At the corner of Wright and Green turned the instructor.
     b. *On his bicycle appeared John.

In the present analysis, these examples are unacceptable simply because an adjunct locative cannot be in the ARG-ST and thus cannot serve as an input to the LI lexicalization.

Weak crossover and Binding: The analysis also presents an account of binding facts.

[4]The mapping from *loc-v-lxm* to *loc-inv-v* could be captured either by a lexical rule or an elaborated type hierarchy system. For the meaning of *can-ss* (*canonical synsem*), see the discussion after (24).

[5]See Manning and Sag (1999) for a similar dissociation between argument structure and valence features in ergative languages.

[6]However, as discussed in Birner (1994), Birner and Ward (1998), and Culicover and Levine (2001), LI allows incorporated nouns in complex verbs., e.g. *take place, take root.*, as in *In that year took place a great renewal.*

(16) a. [Two handsome young boys$_i$] sat [beside each other$_i$].
b. [Beside each other$_i$] sat [two handsome young boys$_i$].

The postverbal NP behaves like the subject since in the LI, the NP binds the anaphor *each other*. Within the binding theory of HPSG where argument-structure is the locus of binding, we then can predict that the 'outranking' NP in (17) can function as the binder of the PP.[7]

(17) $\left[\text{ARG-ST } \langle \boxed{1} \text{ NP[boys}_i\text{]}, \boxed{2} \text{ PP[each other}_i\text{]} \rangle\right]$

Agreement: The subject-verb agreement fact also follows from the assumption that ARG-ST is the locus of the agreement. Since the postverbal NP is the subject at the ARG-ST, the verb will agree with this NP even though it is realized in the postverbal position as a complement. For example, let us see the lexical entry of the singular verb *stands*:

(18) $\begin{bmatrix} \langle \text{stands} \rangle & \\ \text{ARG-ST} & \langle \text{ NP}_{\boxed{i}}, \text{PP}_{\boxed{j}} \rangle \\ \text{CONTENT} & \begin{bmatrix} \text{RELATION} & \textit{stand} \\ \text{THEME} & \boxed{i}\ [\textit{3rd sing}] \\ \text{LOC} & \boxed{j} \end{bmatrix} \end{bmatrix}$

As represented in (19), the verb *stands* requires its theme NP to be 3rd singular. This restriction holds whether the NP is realized as a postverbal complement NP or as subject in a canonical sentence.

Adverb Position: Within the analysis, we also would not expect an adverb to intervene between the verb and the theme NP since the NP is the direct complement of a locative verb. This explains the contrast in (19):

(19) a. Outside the door sat a young man in an attitude of despondency.
b. *Outside the door sat in an attitude of despondency a young man (Penhallurick 1984)

In a similar spirit, this would mean that a heavy object NP in LI could be right-shifted over an adverb as in (20):

(20) a. I read last Tuesday a wonderful book about Korean grammar.
b. On my rug sat on most occasions a magnifi cent waterpipe that I brought back from Nepal.

[7] See Sag and Wasow (1999) for the binding theory of HPSG.

Extraction: In terms of extraction, only the PP can be extracted:

(21) a. *What does in the garden stand _ ?
b. In which garden stands a fountain?

(22) a. *The fountain that in the garden stands _ is my favorite.
b. She stood on the corner on which was standing another woman.

There are even cases where the PP has undergone a long-distance extraction:

(23) a. *What did John say in the garden stood _ ?
b. In which garden did you say _ stood a lamp?

Examples like (23) indicate that we cannot simply specify that LI is a root phenomenon. Why do the NP and the PP in LI behave differently in terms of extraction? We attribute this to the fact that the theme NP in LI cannot be a gapped element: the focused theme NP argument must be a type of canonical syntax-semantics element (*canon-ss*) in order to serve the designated function, the focus of the LI (as represented in (13)).[8] That is, if this NP is realized as a gap element, the LI construction would loose one of its basic functions: making the postverbal NP as representing new information (see section 3.2).

As noted by Bouma et al. (2001), in English there are also peculiar cases where lexemes specify their arguments as being of type of *gap-ss*. One example can be found from the usage of a verb like *assure*:[9]

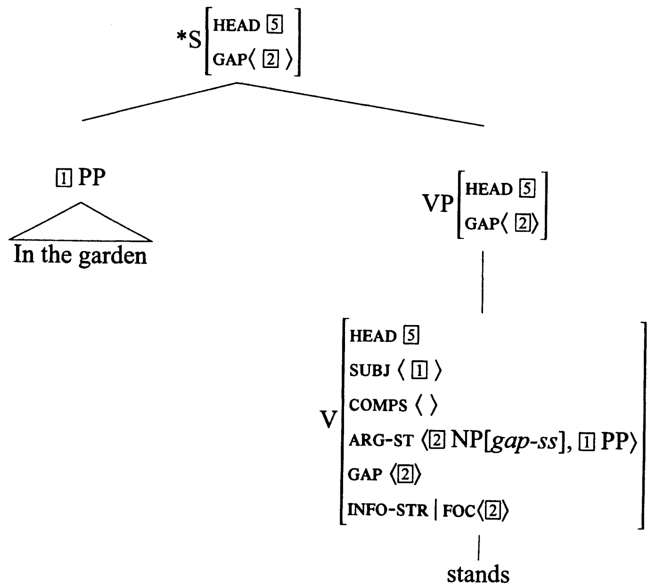
(24) a. This candidate, they assured me _ to be reliable.
b. *They assured me this candidate to be reliable.

What we can observe in (24) is that the understood subject of the infi nitive VP cannot be realized as a canonical NP. Given such a lexical case, it would not unreasonable to assume that the words of *loc-inv-v* impose a specifi c constraint on the type of its focused NP. The lexical specifi cation in (23) means that the NP cannot be realized as a gapped element, blocking examples like (23)a. The structure further explains why this sentence is unacceptable:

[8]Following Bouma et al. (2001), we accept the view that *synsem* has *cannon-ss* and *noncan-ss* as its subtypes, the latter of which in turn has *pro-ss* and *gap-ss*. As default, arguments are *synsem* elements which can be realized either as *cannon-ss* or *noncan-ss*. See Bouma et al. (2001) for details.

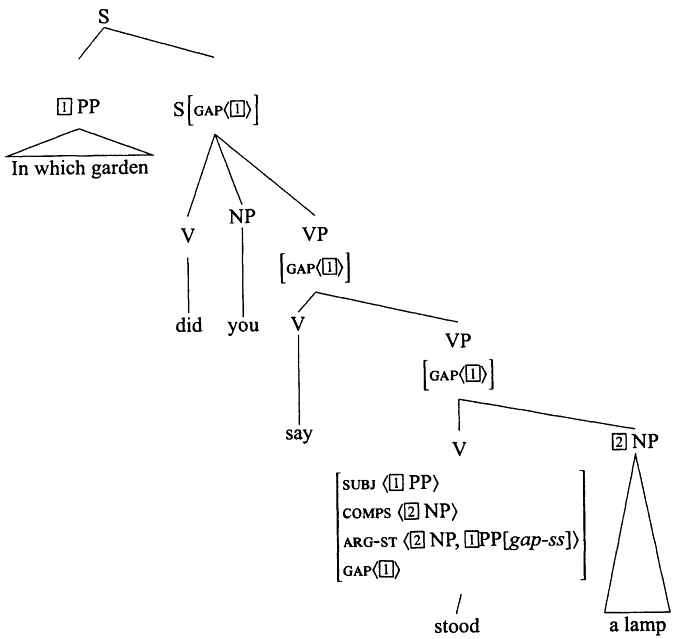
[9]One can argue that this is not about the lexical properties of *assure*, but the general properties of English VP. That is, English allows two NP sisters within VP only when they are the order of Goal-Theme. However, such an ordering restriction is hard to support cases like (24) since thematically *this candidate* is unrelated to *assure* in any way. In addition, there is no problem with *I assure you this candidate is reliable.*

(25)



The present analysis, however, does license cases like (21)b, (22)b, and (23)b. In these examples, what is extracted is not a complement NP but the subject PP. The grammar places no restrictions on the type of this subject PP, implying that it can either be realized as a canonical element or a gapped element. The structural representation of (24) could show us this more clearly:

(26)

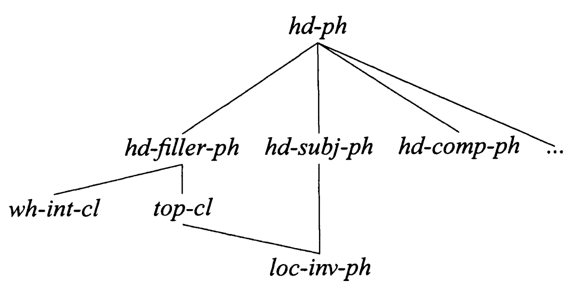


As represented in (27), the subject PP of the *loc-inv-v* is realized as a *gap-synsem* element. This gapped element has been passed up until it is discharged by the fi ller PP *in which garden.* The sentence thus observes all the constraints in question.

3.2. Constructional Constraints

As pointed out earlier, at stake in LI is the fact that the preposed PP has both topic and subject properties, whereas the postposed NP acts like an object and a focus. We attribute these mixed properties to the fact that LI phrases are subtypes of both *hd-subj-ph* and *top-cl*, as represented in a simplifi ed multiple inheritance hierarchy in (27):

(27)



Within a multiple inheritance system, a phrase assigned to a type obtains all the constraints associated with its supertypes, in addition to its own constraints.[10] This then allows us to factor out clausal functions of each phrase while capturing generalizations about particular constructions. The hierarchy in (27) thus makes sure that an instance of *loc-inv-ph* inherits both the constraints of the *topic-cl* (a subtype of *hd-filler-ph*) and those of the *hd-subj-ph* in addition to its own constraint.

A *hd-subj-ph* (e.g., a sentence) and *hd-filler-ph* (like a phrase consisting of [[*Fido*, [*John likes* __]]) independently have the following constraints:

(28) *hd-subj-ph*:

$$\begin{bmatrix}\text{SUBJ} & \langle\ \rangle\end{bmatrix} \rightarrow \boxed{1}\ ,\ \mathbf{H}\begin{bmatrix}\textit{phrase} & \\ \text{SUBJ} & \langle \boxed{1} \rangle\end{bmatrix}$$

(29) *hd-filler-ph*:

$$[\,] \rightarrow \boxed{1},\ \mathbf{H}\begin{bmatrix}\text{HEAD } \textit{verb} \\ \text{GAP} \langle \boxed{1} \rangle\end{bmatrix}$$

(28) says that the *hd-subj-ph* consists of a subject and a VP looking for its subject. Meanwhile, the constraint on *hd-filler-ph* in (29) ensures that the head daughter must be a verb projection and its SLASH value is identifi ed with the LOCAL value of the fi ller daughter. As expected, the *topic-cl* is a subtype of the *hd-filler-ph*, and requires to have at least the following constraints to capture its basic properties:

(30) *topic-cl*:

$$[\,] \rightarrow \boxed{1}\begin{bmatrix}\text{TOPIC} +\end{bmatrix},\ \mathbf{H}\begin{bmatrix}\text{VFORM} \textit{ fin} \\ \text{IC} +\end{bmatrix}$$

According to the constraint in (30), the non-head daughter of a *topic-cl* will carry the topic value, similar to the traditional notion of topic. The constraint also ensures that topicalized clauses are built from independent fi nite clauses, e.g., the declarative head-subject clause, which can serve as independent clauses (as marked by the

[10]The concept of hierarchical classification is essentially assigning phrases (like words) to specific types, and an assignment of those types to superordinate types. Each type is declared to obey certain constraints corresponding to properties shared by all members of that type. This system then allows us to express cross-classifying generalizations about phrases (like words), while accommodating the idiosyncracies of individual types on particular subtypes of words. See Ginzburg and Sag (2001) for a comprehensive study of English interrogative constructions developed within such a multiple inheritance system.

the feature IC (independent clause)).[11] Such basic constraints on topic clauses thus block us from generating sentences like (31):

(31) a. *Bill asked if [such books] John only reads at home.
b. *[Mary tried [the man]$_i$ to kill _ $_i$].

Since clauses headed with *if* or *whether* cannot be independent clauses (thus [IC −]), and thus cannot be topicalized. Also the condition on the fi niteness blocks examples like (31)b. In a similar manner, the constraints in (30) predict the contrast in LI:

(32) a. Mary said [that under the tree sat a woman].
b. *Mary said [under the tree sat Mary].

As shown in Bouma et al. (2001) and Ginzburg and Sag (2001), an embedded clause can be an independent main clause only if the complementizer *that* appears. A *that*-less embedded clause cannot be an independent clause.

One thing to note here is that since *loc-inv-ph* is a subtype of *hd-sub-ph* and *topic-cl* (which is in turn a subtype of *hd-filler-ph*), the phrase inherits the constraints of these supertypes. This in turn means that the phrase will have at least the constraints given in (33):[12]

(33) Constraints on the *loc-inv-ph* which are inherited from its supertypes:

$$[\,] \rightarrow \boxed{1}\begin{bmatrix}\text{TOPIC} & +\end{bmatrix} \; \mathbf{H}\begin{bmatrix}\text{VFORM} & \textit{fin} \\ \text{IC} & + \\ \text{SUBJ} & \langle \boxed{1} \rangle\end{bmatrix}$$

Any instance of *loc-inv-ph* as being a subtype of the two types needs to satisfy the constraints in (33) at least. The only thing we then need to specify on the *loc-inv-ph* is its own constraint as in (34).[13]

(34) *loc-inv-ph*:

$$[\,] \rightarrow \begin{bmatrix}\text{HEAD } \textit{prep} \\ \text{LOCATION } i\end{bmatrix}, \mathbf{H}\begin{bmatrix}\text{INFO-ST} \mid \text{FOC } \textit{nelist}\end{bmatrix}$$

[11]Thus cases like relative clauses cannot have a topic nonhead daughter. See Ginzburg and Sag (2001).

[12]The phrase *loc-inv-ph* is specified not to inherit from *hd-fi ller-ph* the constraint that the topic phrase is the filler of the gap value of the head.

[13]Following Engdahl and Vallduví (1996), I assume that the FOCUS value of a daughter is inherited up to its mother.

What the constraint in (34) says is that a *loc-inv-ph* consists of a locative phrase as non-head daughter and a head element which has a focus value. All the other constraints, such as the assignment of topic value to the PP, will be inherited from its supertypes.

By cross-classifying *loc-inv-ph* as a subtype of *topic-cl* and *hd-subj-ph*, we could account for why *loc-inv-ph* has various filler properties. For example, as pointed out by Bresnan (1994), raising asymmetries in (35) could be another indicator for the filler properties of the preverbal PP.

(35) a. *I [[$_{VP}$ expect] [$_{PP}$ on this wall] [$_{VP}$ to be hung a portrait of our founder]].
  b. [On this wall] [is likely to be [hung a portrait of our founder]].

Within our system, (35)a violates two constraints at least. First of all, the PP and the VP does not form a constituent as given by the bracket, thus not forming a *loc-inv-ph*. Another violation relates to the topichood of the locative PP. Since it carries a topic value, a locative phrase should serve as the nonhead daughter or a filler of a finite clause. This is not the case in (35)a.[14]

Such a construction view of grammar also provides an answer to the lack of Yes-no question:

(36) a. *Did [into the room] [walk a woman]?
  b. *[Into the room] I expected _ to [walk Robin].

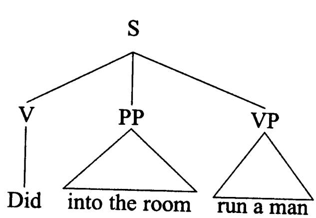
Adopting the idea of Fillmore (1999) and Ginzburg and Sag (2001), we accept the view that English has the construction of *sai-ph* whose constraints are given in (37):

(37) *sai-ph*:

$$\left[\text{SUBJ}\ \langle\quad\rangle\right] \rightarrow \mathbf{H}\begin{bmatrix}\text{INV} + \\ \text{AUX} + \\ \text{SUBJ}\ \langle \boxed{1} \rangle \\ \text{COMPS}\ \boxed{A}\end{bmatrix}, \boxed{1}, \boxed{A}$$

What this means is that a *sai-ph* consists of an auxiliary head, its subject ($\boxed{1}$) and complement(s) ($\boxed{A}$). This will then assign the following structure to a yes-no sentence like (36)a:

(38)



[14]If this forms a small clause, it violates the finiteness condition of a topic clause.

The S structure in (38), however, is not a *loc-inv-ph* since the topic PP *into the room* does not form a phrase (constituent) with the head VP *run a man*. As given in the *loc-inv-ph* constraints in (33) and (34), a topic phrase can occur only as the non-head daughter with a fi nite head phrase. In (38), the VP headed by the locative verb *run* is neither fi nite nor the head: the head is the auxiliary verb *did*. The present grammar thus would not generate an yes-no LI sentence.

4. Conclusion

In the English LI construction, the preverbal PP and postverbal NP have mixed functional properties. This paper presents a simple way of dealing with mixed functional properties in HPSG. This paper is very similar to the analysis of Schachter (1992) and Bresnan (1994) in that the PP in LI is a topicalized subject whereas the NP is a focus.

However, this analysis differs from such previous analyses in several respects. In syntax the preverbal PP is not in the topic position, but it is in the subject position. In addition, the postverbal NP is a direct complement of a locative verb, not adjoined to the VP. The analysis also accept the existence of *loc-inv-ph* in English that is a subtype of both *topic-cl* and *hd-sub-ph*. We have shown that the mixed, intricate properties are due to the interaction between the lexical information of the verbs in LI and its constructional constraints.

References

Birner, Betty J. 1994. Information Status and Word Order: An Analysis of English Inversion. *Language* 70: 233-259.

Birner, Betty J. and Gregory Ward. 1998. *Information Status and noncanonical Word Order in English.* Amsterdam/Philadelphia: John Benjamins.

Bouma, Gosse, Robert Malouf and Ivan Sag. 2001. Satisfying constraints on extraction and Adjunction. *Natural Language and Linguistic Theory* 19, 1-65.

Bowers, John. 1976. On surface structure grammatical relations and the structure-preserving hypothesis. *Linguistic Analysis* 2, 225-242

Bresnan, Joan. 1994. Locative Inversion and the Architecture of Universal Grammar, *Language* 70, 72-131.

Bresnan, Joan and Jonni Kanerva. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20, 1-50.

Chung, Chan and Jong-Bok Kim. 2002. Lexical and Constructional Constraints in English Locative Inversion. Paper presented at the 38th Meeting of the Chicago Linguistic Society The University of Chicago, April 25–27, 2002.

Culicover, Peter and Robert Levine. 2001. Stylistic inversion in English: a reconsideration. *Natural Language and Linguistic Theory* 19, 283-310.

Coopmans, Peter. 1989. Where stylistic and syntactic processes meet: Locative inversion in English. *Language* 65, 728-751.

Engdahl, Elizabeth and Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover and Enric Vallduvi (eds.), *Studies in HPSG*: Edinburgh Working Papers in Cognitive Science, 1–31.

Fillmore, Charles. 1999. Inversion and Constructional Inheritance. In *Lexical and Constructional Aspects of Linguistic Explanation,* 113-128. Stanford: CSLI Publications.

Green, Georgia. 1992. Towards an HPSG Account of So-called 'Focus Inversion', Paper Presented at HPSG Workshop at Stanford Univ.

Ginzburg, Jonathan and Ivan Sag. 2001. *Interrogative Investigations* CSLI Publications, Stanford.

Hoekstra, Teun, and Ren Mulder. 1990. Unergatives as copular verbs: Locational and Existential predication. *The linguistic review* 7, 1-79.

Kathol, Andreas and Robert Levine. 1992. Inversion as a linearization effect. *Proceedings of North East Linguistic Society 23, North East Linguistic Society,* pp. 207-221.

Levine, Robert. 1989. On focus inversion: syntactic valence and the role of a SUBCAT list. *Linguistics* 17. 1013-1055.

Levine, Robert. 2002. Paper Presented at the 2002 LSK International Conference on Linguistics.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-lexical Semantics Interface.* Cambridge, Mass: MIT Press.

Manning Chris and Ivan Sag. 1999. Dissociations between ARG-ST and grammatical relations. *Lexical and Constructional Aspects of Linguistic Explanation.* CSLI Publications, Stanford, pp. 63-78.

Penhallurick, John. 1984. *Full verb inversion in English. Australian Journal of Linguistics*, 4: 33–56.

Rochemont, Michael. 1986. *Focus in generative grammar.* Amsterdam: Benjamins.

Rochemont, Michael and Peter Culicover. 1990. *English Focus Constructions and the Theory of Grammar.* Cambridge University Press, Cambridge: Cambridge University Press.

Sag, Ivan and Tom Wasow. 1999. *Syntactic Theory: A Formal Introduction.* Stanford: CSLI Publications.

Schachter, Paul. 1992. Comments on Bresnan and Kanerva's "Locative Inversion in Chichewa: A Case Study of Factorization in Grammar." In Tim Stowell and

Eric Wehrli (eds.), *Syntax and Semantics 26: Syntax and the Lexicon*, 103-116.

Jong-Bok Kim
School of English
Kyung Hee University, Seoul
jongbok@khu.ac.kr

Contrast and Phonological Activity in the Nez Perce Vowel System*

SARA MACKENZIE and B. ELAN DRESHER
University of Toronto

0. Introduction

Features that pattern as inert with respect to phonological processes are often unnecessary in distinguishing segments in an inventory. A natural way of accounting for the phonological inertness of noncontrastive features is via underspecification. If noncontrastive features are absent from underlying representations and barred from the phonology, it follows that they will not be visible to phonological processes. However, there has been no consensus on how to determine which features are contrastive in any given case. This and other objections have detracted from the appeal of theories of underspecification.

In Optimality Theory (OT, Prince and Smolensky 1993), moreover, representational devices such as underspecification and constraints on underlying forms have been generally rejected in favour of constraints on output forms. Inventories arise from constraint ranking, and the role of the inventory in determining phonological behaviour is therefore quite limited. However, the question of how to account for the relationship between contrast and phonological activity within OT remains open.

The theory of the contrastive hierarchy (Jakobson and Halle 1956, Dresher 1998b, 2003) provides a method for determining contrasts in an inventory by ordering features so that some features take scope over others. This theory overcomes the major objections to underspecification theory, and connects contrast to phonological activity, because the size and shape of an inventory will affect the number of features needed in order for each segment to be contrastively specified.

This paper provides an analysis of the Nez Perce vowel system within the framework of the contrastive hierarchy. Our analysis follows Jacobsen (1968), Rigsby and Silverstein (1969), Zwicky (1971), and Hall and Hall (1980) in assuming an abstract front vowel that participates in ATR harmony. In addition, we demonstrate that contrastive specifications are compatible with OT and can be achieved using mechanisms which are central to that theory, such as IDENT constraints and contextual markedness constraints.

* Some parts of this paper are based on Mackenzie 2002. We would like to thank members of the project on Contrast in Phonology at the University of Toronto for helpful discussion (http://www.chass.utoronto.ca/~contrast/). This research was supported in part by grants 410-96-0842 and 410-99-1309 from the Social Sciences and Humanities Research Council of Canada.

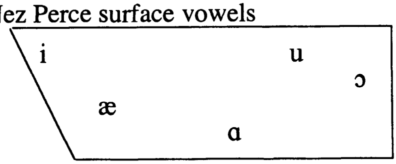
An alternative analysis of Nez Perce vowel harmony within the framework of OT is proposed by Bakovic (2000). He uses featural faithfulness constraints and markedness constraints. His account differs from ours in its lack of reference to contrast. The case is of some interest, because in the analysis of Bakovic (2000) the feature [ATR], which is phonologically active in the vowel system of Nez Perce, appears to be noncontrastive, contrary to the observation we made at the outset that noncontrastive features tend to be phonologically inert.

Appealing to the theory of the contrastive hierarchy, we will demonstrate that the crucial constraint rankings in Bakovic 2000 are unmotivated, and that [ATR] is a contrastive feature in the Nez Perce vowel system. Whereas every contrastive hierarchy can be expressed as a set of ranked OT constraints, not every OT constraint hierarchy can be converted to a contrastive hierarchy. Requiring grammars to adhere to the contrastive hierarchy can help to determine the crucial constraint rankings and constrain the set of possible grammars.

1. The Nez Perce Vowel System (Aoki 1966, 1970, Hall and Hall 1980)

The surface vowels of Nez Perce are shown in (1) (Aoki 1966, 1970):

(1) Nez Perce surface vowels



Nez Perce has dominant-recessive ATR harmony. All vowels in a word apart from /i/ must agree with respect to [ATR], and the value [–ATR] is dominant. /æ/ alternates with /ɑ/ (2) and /u/ alternates with /ɔ/ (3).

(2) ATR harmony: /æ/ alternates with /ɑ/

	Underlying	*Surface*	*Gloss*
a)	/næʔ-mæq/	næʔmæχ	'my paternal uncle'
b)	/næʔ-tɔt/	nɑʔtɔt	'my father'
c)	/mæq-æʔ/	mæqæʔ	'uncle VOC'
d)	/tɔt-æʔ/	tɔ:tɑʔ	'father VOC'
e)	/cæqæt/	cæ:qæt	'raspberry'
f)	/cæqæt-ɑyn/	cɑqɑ:tɑyn	'for a raspberry'

(3) ATR harmony: /u/ alternates with /ɔ/

	Underlying	*Surface*	*Gloss*
a)	/tæwæ:-pu:/	tæwæ:pu:	'the people of Orofino, Idaho'
b)	/sɔ:yɑ:-pu:/	sɔ:yɑ:pɔ:	'the white people'
c)	/tuʔuynu/	tuʔuynu	'tail'
d)	/tuʔuynu-ʔɑyn/	tɔʔɔynɔʔɑyn	'for the tail, crupper'

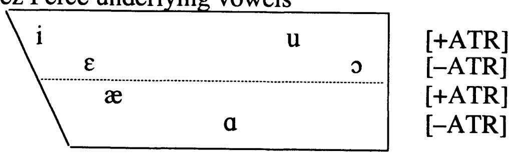
As illustrated in (4), the vowel /i/ sometimes patterns with [–ATR] vowels (4a,b), and other times with [+ATR] vowels (4c,d), though it is phonetically [+ATR].

(4) Dual patterning of /i/

	Underlying	*Surface*	*Gloss*
a)	/næʔci:c/	nɑʔci:c	'my paternal aunt'
b)	/ ci:c-æʔ/	ci:cɑʔ	'paternal aunt VOC'
c)	/næʔ-i:c/	næʔi:c	'my mother'
d)	/ʔi:c-æʔ/	ʔi:cæʔ	'mother VOC'

Following Hall and Hall (1980) we assume that surface [i] represents a merger of /i/ and a [–ATR] vowel we will represent as /ɛ/. In (4a, b) the underlying vowel is /ɛ/ and in (4c, d) the vowel is [+ATR] /i/.

(5) Nez Perce underlying vowels



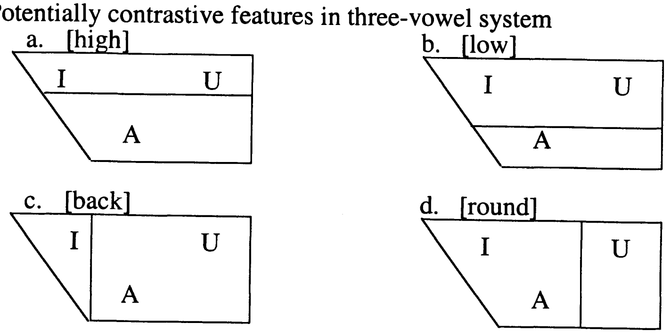
Thus, every vowel has a counterpart that contrasts with it in the feature [ATR].

2. Contrastive specification by a hierarchy of features

By any definition, [ATR] is a *contrastive* feature in the underlying vowel system of Nez Perce. What are the other contrastive features? Abstracting away from [ATR], we have a classic three-vowel system, which we can designate /I A U/. Even in such a simple system, it is not obvious what the relevant contrasts are. We need a way to determine contrasts in an inventory.

The approach we will take is that of Jakobson and Halle (1956). To determine contrastiveness of features, it is necessary to determine their relative *scope*, or *ordering*. In a simple three-vowel system, for example, exactly two features are needed for each segment to be contrastively specified, though what these are can potentially vary. Some candidates are shown below:

(6) Potentially contrastive features in three-vowel system



Ordering is required to select contrastive features. It also determines the relative scopes of the contrastive features that are selected. Suppose, for example, we choose the features [high] and [round]. We can first divide the vowels on the basis of [high] (7a). Now [round] is relevant only as a contrast among the [+high] vowels: /i/ and /u/ are 'partners', /a/ is neutral. In this ordering, [high] > [round], all vowels are contrastive for [high], but only /i/ and /u/ are contrastive for [round]. Alternatively, we can first divide the vowels on the basis of [round] (7b). Now [high] is relevant only among the [–round] vowels: /i/ and /a/ are 'partners', /u/ is neutral. In this ordering, [round] > [high], all vowels are contrastive for [round], but only /i/ and /a/ are contrastive for [high].

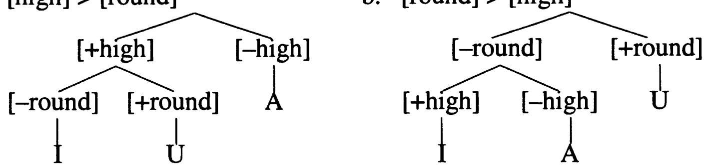
(7) Scope or ordering of features

a. [high] > [round] b. [round] > [high]

These two orderings correspond to two different contrastive hierarchies, as shown in (8). Given a contrast based on [high], a further contrast based on [round] is required only for members that are specified as [+high] (8a). Given a contrast based on [round], a further contrast based on [high] is required only for members that are specified as [–round] (8b).

(8) Scope or ordering of features

a. [high] > [round] b. [round] > [high]

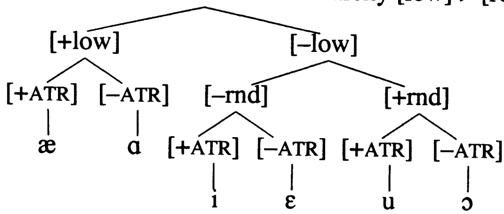


We might expect that the two vowel systems will pattern differently. For example, system (8a) might show alternations or neutralization between /i/ and /u/; in system (8b) /i/ might be more closely related to /a/.

We assume that the contrastive hierarchy can vary within limits determined by Universal Grammar (UG). In determining what the hierarchy is for any particular language it is necessary to examine phonological processes for evidence as to what the active features are.

In the case of Nez Perce, we choose [low] as the first feature in the hierarchy. We choose [low] rather than [high] because the surface nonlow [ATR] pair is not strictly [+high], whereas the low pair are both [+low]. The choice of [low] as the initial feature is also consistent with Jakobson and Halle's (1956) assumption that a contrast between high and low sonority is, preferably, ordered before one based on place (but see Ghini 2001 for a different view). We choose [round] as the second feature. This contrast is relevant only among the nonlow vowels. Because of the symmetry of the system, it does not matter very much where [ATR] is ordered. For concreteness, we will assume it is ordered third. We thus arrive at the contrastive hierarchy illustrated in (9).

(9) Nez Perce vowels: Contrastive hierarchy [low] > [round] > [ATR]



The above approach to contrastive specification by a hierarchy of features can be implemented by an algorithm called the Successive Division Algorithm (SDA) (Dresher 1998a, 1998b, 2002, 2003, Dresher and Zhang this volume). An informal version is given in (10).

(10) Successive Division Algorithm (SDA)
- a. Begin with *no* feature specifications: assume all sounds are allophones of a single undifferentiated phoneme.
- b. If the primordial allophonic soup is found to consist of more than one contrasting member, select a feature and divide the set into as many subsets as the feature allows for.
- c. Repeat step (b) in each subset: keep dividing up the inventory into sets, applying successive features in turn, until every set has only one member.

3. The Contrastive Hierarchy and Specification

The contrastive hierarchy is not contingent on a particular theory of phonological operations or representations. In particular, it does not necessarily presuppose underspecification. At its most basic, the SDA assigns contrastive features. What about redundant features?

One possibility is that all features are specified, and the algorithm designates some as contrastive. Certain phonological processes can then be designated as targeting only contrastive features (Calabrese 1995). A stronger theory would be one that makes redundant features unavailable to the (lexical) phonology except under special conditions. Such a restriction is captured in a natural way by supposing that only features assigned by the SDA are specified.

In recent years a number of arguments have been raised against underspecification that have appeared to undermine such an approach. However, the contrastive hierarchy puts these issues in a new light.

First, it has been argued (Steriade 1995, Kirchner 1997) that there is no consistent way to decide which specifications to omit. For example, in most languages there are no nasal obstruents, so [+sonorant] is predictable given [+nasal]; nevertheless, this specification is rarely omitted. The conclusion is that underspecification is applied inconsistently. The answer to this objection is that the contrastive hierarchy decides which features are omitted. In particular, phonological redundancy is not the same as logical redundancy. Many features that are logically redundant are designated as contrastive by the SDA. Indeed, the arbitrariness argument applies not only to underspecification, but also to contrast. It is necessary to have a principled method for determining which features are contrastive as well as which features may be unspecified. The contrastive hierarchy and the SDA provide such a method.

In the above example, it is more common for [sonorant] to take scope over [nasal] than it is for [nasal] to take scope over [sonorant]. Therefore, [+sonorant] must be specified even where it is made logically redundant by [+nasal]. The hierarchy [nasal] > [sonorant] is less likely and leads to an unusual set of contrasts in an inventory.

Second, it has been argued that there is relatively little evidence for underspecification. This argument assumes that full specification is the null hypothesis, unless positive evidence is found to the contrary. Thus, the burden of proof has been placed on underspecification. But it is not clear that the burden of proof *should* be on underspecification. There is in fact no positive evidence for full specification. In practice, most analyses that reject underspecification do not adopt full specification: features totally irrelevant to an analysis are rarely specified. The result is not full specification but arbitrary specification.

Some recent approaches start from the premise that features are specified only if there is positive evidence to do so. Examples are Modified Contrastive Specification as developed in Toronto (Avery and Rice 1989, Avery 1996, Ghini 2001, Rice 2002, Hall 2003); the theory of representational economy of Clements (2001); and the system-driven specification of Hyman 2002a, 2002b.

4. The Contrastive Hierarchy in Optimality Theory

Some ways of implementing the contrastive hierarchy are inconsistent with assumptions commonly made in OT to the effect that (a) there is no necessary underspecification (Itô, Mester and Padgett 1995), and (b) there are no limitations on underlying inventories (Prince and Smolensky 1993). These issues concern implementation of the contrastive hierarchy, and are orthogonal to the notion of the contrastive hierarchy of features itself. We will show that the contrastive hierarchy can be modeled in OT.

It has been claimed (Kirchner 1997) that contrasts emerge from constraint rankings, so one might think that there is no need to say anything more about contrast. But an arbitrary constraint ranking will not express a connection between contrast and phonological activity. If there is such a connection, it should be captured in any phonological theory. OT analyses that are consistent with a contrastive hierarchy can express this connection.

In converting the contrastive hierarchy into an OT constraint set, we must make some assumptions about the output and the input.

Output: We will assume that the output of an OT version of the SDA is the same as the output of the algorithm itself: a set of contrastive specifications from which redundant feature specifications are excluded. In the case of Nez Perce, we will also assume that the output of this evaluation contains the [–ATR] counterpart of /i/. We will not attempt to model [ATR] harmony in this algorithm. How the processes of neutralization and [ATR] harmony) are to be incorporated is not crucial to our proposal for modeling contrast.

Input: We will assume for now that the input consists of fully-specified representations. The analysis can be extended to include underspecified inputs, but we shall not do so here.

Constraints: We will employ two basic constraint types, feature-specific IDENT constraints and contextual markedness constraints of the form $*[\alpha F, \Phi]$.

(11) Constraints regulating contrast

a. IO-IDENT F: 'Correspondent segments must have the same value of the feature F (either + or –)'.

b. *[αF, Φ]: 'Exclude αF in the context Φ', where α ranges over + and –, and Φ is the set of features (with wider scope than F) forming the context of F.

To convert a contrastive hierarchy into a constraint ranking, faithfulness constraints referring to contrastive features will be ranked in the same order as the features in the contrastive hierarchy. At each point when a segment is uniquely specified a contextual markedness constraint will be introduced, preventing further specification.

In the case of Nez Perce the first feature in the hierarchy is [low]. It has no exclusions. Therefore, we place the constraint IO-IDENT [low] in the highest constraint stratum, ensuring that the underlying value of this feature is maintained.

The second feature is [round]. It is excluded with [+low], because the feature [round] does not serve to further distinguish segments in the [+low] set. Hence the constraint *[+low, round] must be ranked ahead of IO-IDENT [round]. Thus, no segment with underlying value [+low] will be able to surface with a value for [round]; [–low] segments must keep their underlying value of [round].

The third feature is [ATR]. It also has no exclusions, so we place the constraint IO-IDENT [ATR] next in the constraint hierarchy.

All other vowel features are redundant and are excluded. We can obtain this result by adding the constraint *[F], which filters out all but contrastive specifications that have higher-ranking IO-IDENT constraints. The resulting constraint hierarchy is summarized in (12).

(12) Nez Perce constraint hierarchy regulating contrasts in the inventory
IO-IDENT [low] >> *[+low, round] >> IO-IDENT [round] >> IO-IDENT [ATR] >> *[F]

A general algorithm for converting a contrastive hierarchy to an OT constraint hierarchy, given an ordering of features, is shown in (13).

(13) Converting a contrastive hierarchy into a constraint hierarchy

a. Select the next contrastive feature in the list, F_i. If there are no more contrastive features, go to (e).
b. In the next stratum, place any co-occurrence constraints of the form *[αF_i, Φ], where Φ consists of features ordered higher than F_i.
c. In the next stratum, place the constraint IO-Ident [F_i].
d. Go to (a).
e. In the next constraint stratum, place the constraint *[F], and end.

Every contrastive hierarchy can be converted into a constraint hierarchy by the above procedure. But the converse does not hold: not every constraint hierarchy corresponds to a well-formed contrastive hierarchy.

5. The Analysis of Bakovic 2000

An OT analysis of the Nez Perce vowel system is given in Bakovic 2000. His analysis has some properties in common with ours. Like us, he needs a hierarchy of featural faithfulness constraints, and constraints to exclude certain combinations of features. However, he goes about defining the Nez Perce vowel inventory in quite a different fashion.

Bakovic begins with the following table showing the actual Nez Perce surface vowels (in bold) and a number of nonexistent vowels that need to be excluded:

(14) Features of existing and absent vowels (Bakovic 2000:243 (270))

	[+ATR]		[−ATR]	
[+high, −low]	i	u	ɪ	ʊ
[−high, −low]	e	o	ɛ	ɔ
[−high, +low]	æ		e	ɑ
	[−back]	[+back]	[−back]	[+back]

Starting from the assumption that inputs are not restricted to language-specific inventories, Bakovic introduces constraints to derive the surface inventory. To ensure faithful mapping of the actual vowels, he employs feature-specific IO-IDENT constraints. To penalize absent vowels, he uses markedness constraints that militate against antagonistic tongue gestures. For example, the constraint *[−high, +ATR] suggests there is an antagonism between a lowered tongue body and an advanced tongue root.

The IDENT constraints refer to the features shown in (14). But no arguments are given for why these features were selected and others omitted. This is neither full nor contrastive specification, but rather arbitrary specification. Viewed against our contrastive hierarchy for Nez Perce, the features [high] and [back] are redundant, and a contrastive feature, [round], is missing.

According to Bakovic, the motivation for a high ranking of faithfulness to [high] is to ensure that a hypothetical input /o/, a vowel which does not exist in Nez Perce, will surface as [ɔ] rather than as [u] (15). But no evidence is adduced that an input /o/ does in fact surface as [ɔ] and not, say, as [u]. Therefore, the relatively high ranking of this constraint has no real motivation, for we cannot exclude a constraint hierarchy such as in (16).

(15) Role of IO-IDENT[hi] (Bakovic 2000:245 (273))

Input /o/	*[−high, +ATR]	IO-IDENT [high]	IO-IDENT [ATR]
a. o	*!		
b. u		*!	
c. ☞ ɔ			*

(16) Evaluation of /o/ with low-ranking IO-IDENT[hi]

Input /o/	*[−high, +ATR]	IO-IDENT [ATR]	IO-IDENT [high]
a. o	*!		
b. ☞ u			*
c. ɔ		*!	

Similarly, Bakovic (2000:246) wishes to ensure that input /e/ surfaces as [i]. In his analysis, faithfulness to [back] plays a prominent role in preventing /e/ from surfacing as *[ɔ] (17). Again, there are many other ways of excluding this vowel; one is shown in (18).

(17) Role of IO-IDENT[bk] (adapted from Bakovic 2000:246 (279))

Input /e/	IO-ID [low]	*[–bk, –ATR]	IO-ID [back]	*[-hi, +ATR]	IO-ID [high]	IO-ID [ATR]
a. e				*!		
b. ☞ i					*	
c. ɛ		*!				*
d. ɔ			*!		*	*
e. æ	*!			*	*	

(18) Evaluation of /e/ without IO-IDENT[bk]

Input /e/	IO-ID [low]	*[+low, round]	IO-ID [round]	IO-ID [ATR]	*[F]
a. e					*! [hi]
b. ☞ i					
c. ɛ				*!	
d. ɔ			*!	*	
e. æ	*!				*

Proceeding in this way, Bakovic arrives at the ranking shown in (19).

(19) Constraint ranking for Nez Perce (Bakovic 2000)
*[+back, +ATR] & IO-ID [ATR], IO-ID [low], *[–back, –ATR] >> IO-ID [bk] >> *[–high, +ATR], *[+high, –ATR] >> IO-ID [high] >> *[+back, +ATR] >> *[-low, –ATR] >> IO-ID [ATR]

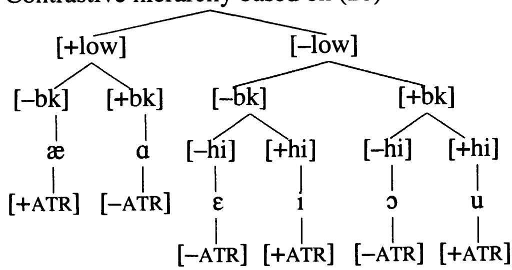
He proposes that these faithfulness constraints and cooccurrence restrictions are sufficient to exclude nonexistent vowels and to ensure that vowels present in the inventory will surface faithfully.

Looking only at the faithfulness constraints, we find the hierarchy in (20).

(20) Ranking of faithfulness constraints (Bakovic 2000)
IO-ID [low] >> IO-ID [back] >> IO-ID [high] >> IO-ID [ATR]

This constraint hierarchy translates into an ill-formed contrastive hierarchy:

(21) Contrastive hierarchy based on (20)



The feature [ATR] is redundant in this hierarchy, though it is the active feature in vowel harmony. It is redundant because of the presence of [high], which does not appear in our contrastive hierarchy. Recall that the motivation for a high ranking of faithfulness to [high] is to ensure that input /o/ will surface as [ɔ] rather than [u] (15). But an analysis that adheres to the contrastive hierarchy automatically prevents illicit vowels from surfacing. In this case, an /o/ has the same contrastive features as /u/; no other features may be specified (16).

The analysis in Bakovic 2000 thus appears to require a ranking of faithfulness constraints that is incompatible with any contrastive hierarchy for Nez Perce. Moreover, this analysis does not draw any connection between contrastiveness and phonological activity in Nez Perce. Given its low ranking, the feature [ATR] appears to be redundant, though it is the active feature in vowel harmony. However, this ranking is unmotivated by any empirical facts and relies primarily on unsupported assumptions about what nonexistent vowels should map to.

6. Conclusions

We have shown that the contrastive hierarchy can be implemented in OT using the same sort of constraints already in common use in the theory. We have also argued that the Nez Perce vowel system can be analyzed using only contrastive features. Although there are a number of possible contrastive hierarchies for Nez Perce, the analysis of Bakovic 2000 does not correspond to any of them. We conclude in particular that the feature [ATR] is contrastive in Nez Perce; thus, vowel harmony is implemented by a contrastive feature.

Finally, limiting OT constraint hierarchies to those that correspond to contrastive hierarchies provides a principled method for choosing between a number of possible constraint rankings. This limitation represents a significant restriction on the set of possible grammars.

References

Aoki, Haruo. 1966. Nez Perce vowel harmony and Proto-Sahaptian vowels. *Language* 42: 759-767.

Aoki, Haruo. 1970. *Nez Perce Grammar.* Berkeley, CA: University of California Publications in Linguistics.

Avery, Peter. 1996. The representation of voicing contrasts. Ph.D. dissertation. University of Toronto.

Avery, Peter and Keren Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6: 179-200.

Bakovic, Eric. 2000. Harmony, dominance and control. Ph.D. Dissertation. Rutgers University.

Calabrese, Andrea 1995. A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26: 373-463.

Clements, G. N. 2001. Representational economy in constraint-based phonology. In T. Alan Hall (ed.) *Distinctive Feature Theory*, 71-146. Berlin: Mouton de Gruyter.

Dresher, B. Elan. 1998a. Child phonology, learnability, and phonological theory. In Tej Bhatia and William C. Ritchie (eds.) *Handbook of Language Acquisition*, 299-346. New York: Academic Press.

Dresher, B. Elan. 1998b. On contrast and redundancy. Paper presented at the annual meeting of the Canadian Linguistic Association, Ottawa. Ms., University of Toronto.

Dresher, B. Elan. 2002. The contrastive hierarchy in phonology. Paper presented at the Second International Conference on Contrast in Phonology, Toronto. In Hall (2003).

Dresher, B. Elan. 2003. Contrast and asymmetries in inventories. In Anna-Maria di Sciullo (ed.) *Asymmetry in Grammar, Volume 2: Morphology, Phonology, Acquisition*, 239-257. Amsterdam: John Benjamins.

Dresher, B. Elan and Xi Zhang. This volume. Phonological contrast and phonetics in Manchu vowel systems.

Ghini, Mirco. 2001. Place of articulation first. In T. Alan Hall (ed.) *Distinctive Feature Theory*, 147-176. Berlin: Mouton de Gruyter.

Hall, Beatrice L. and R. M. R. Hall. 1980. Nez Perce vowel harmony: An Africanist explanation and some theoretical consequences. In R. M. Vago (ed.) *Issues in Vowel Harmony*, 201-236. Amsterdam: John Benjamins.

Hall, Daniel Currie (ed.). 2003. *Toronto Working Papers in Linguistics* 20 (*Special Issue on Contrast in Phonology*). Toronto: Department of Linguistics, University of Toronto.

Hyman, Larry M. 2002a. 'Abstract' vowel harmony in Kàlɔ̀ŋ: A system-driven account. Paper presented at Théories Linguistiques et Langues Sub-Sahariennes, Paris VIII (St Denis). Ms., University of California, Berkeley.

Hyman, Larry M. 2002b. Vowel harmony in Gunu. Paper presented at the 3èmes Journées Internationales du GDR 'Phonologie', Nantes, May 2001. Ms., University of California, Berkeley.

Itô, Junko, Armin Mester and Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26: 571-613.

Jacobsen, William H., Jr. 1968. On the prehistory of Nez Perce vowel harmony. *Language* 44: 819-829.

Jakobson, Roman and Morris Halle. 1956. *Fundamentals of Language*. The Hague: Mouton.

Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14: 83-111.

Mackenzie, Sara. 2002. Contrast, inventories, and vowel harmony in OT. Paper presented at the annual meeting of the Canadian Linguistic Association, Toronto. Ms., University of Toronto.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, N.J.

Rice, Keren 2002. Vowel place contrasts. In Mengistu Amberber and Peter Collins (eds.) *Language Universals and Variation*, 239-270. Westport, CT: Praeger.

Rigsby, Bruce J. and Michael Silverstein. 1969. Nez Perce vowels and Proto-Sahaptian vowel harmony. *Language* 45: 45-59.
Steriade, Donca. 1995. Underspecification and markedness. In John A. Goldsmith (ed.) *Handbook of Phonology*, 114-174. Oxford: Blackwell.
Zwicky, Arnold M. 1971. More on Nez Perce: An alternative analysis. *International Journal of American Linguistics* 37: 122-126.

Department of Linguistics
University of Toronto
Toronto, Ontario
Canada M5S 3H1

s.mackenzie@utoronto.ca
dresher@chass.utoronto.ca

On morphological case and word-order freedom[1]

THOMAS MCFADDEN
University of Pennsylvania

1. The correlation and two ways to explain it

A commonly made cross-linguistic generalization is that languages with extensive case-marking tend also to have greater freedom of word order than languages without. Thus case-marking languages like Latin allow the order of nominal arguments in a sentence to be permuted without changing the basic meaning of the sentence, (discourse-pragmatic factors aside) as in (1) and (2), while swapping the order of arguments in an essentially caseless language like English is not possible without swapping their grammatical roles, as in (3) and (4):

(1) Puella puerum videt.
girl:NOM boy:ACC sees
`The girl sees the boy.'

(2) Puerum puella videt.
boy:ACC girl:NOM sees
`The girl sees the boy.'

(3) The girl sees the boy.

(4) The boy sees the girl.

That a correlation of this sort exists is not really a matter of debate. What is controversial is what form it takes and exactly how we are to explain it.

I will start my discussion by laying out some basic assumptions. First of all, I assume that there is a core of formal linguistic knowledge or grammar which is distinct from the knowledge relating to language use. I further assume that a complete linguistic theory should incorporate tools to deal with both sides of this distinction and provide a principled way to determine which tools should be used to explain a given phenomenon. Thus, in place of the loaded terms *competence*

---

[1] Thanks to Tony Kroch, Dave Embick, Beatrice Santorini and the audience at BLS 29 for their comments and criticisms, and to the following native-speaker informants for their judgments and discussion: Beatrice Santorini, Augustin Speyer, Hans Van de Velde, Hedde Zeijlstra, Sophia Malamud, Anna Nagayeva and Kimiko Nakanishi.

and *performance*, I will distinguish *(narrow) synchronic grammar* from the more differentiated *(principles of) use, acquisition and change*. While a full theory of this type is still a good ways off, the following are reasonable as guidelines. Principles of the synchronic grammar deal with derivation and representation and are stated in formal/structural terms, while principles of use, acquisition and change make reference to the speaker/hearer, to constraints on ease of production, comprehension and acquisition, to social factors, and to the effects of context.

It is clear that the case-marking/word-order freedom correlation (henceforth CWC) could be explained in terms of use, acquisition and change. Namely, word-order variation is easier to interpret, learn, and maintain over time in a language which marks grammatical relations with m-case. However, we can also imagine that this constraint is encoded directly within the narrow synchronic grammar, that the principles of grammar which determine word order make reference to m-case, such that rich case-marking licenses word orders which are impossible in its absence. A number of explanations of this latter type have been proposed over the years, of which I will mention three recent examples here. Roberts (1997), assuming a Minimalist framework, argues that the strength of Case features which drive DP movement is directly related to overt case-marking. Kiparsky (1997) develops a system in which the association of a DP with its θ-role must be guided either by overt case-marking or by canonical ordering. A DP can thus only leave its canonical position if case-marked. Finally, Neeleman and Weerman (1999) propose that Case is a syntactic head K which, when null, as in English, is forced by the ECP to remain in a position properly governed by its licensing head. An overt K is not subject to ECP, thus is free to move.

Now, it's clear that the pressures placed on the speaker-hearer as described above play at least **some** role in the restrictions on word-order freedom. Thus the burden of proof is arguably on those like Roberts, Kiparsky and Neeleman and Weerman who propose an additional synchronic-grammatical component to the correlation. They must show that such a hybrid explanation can account for facts that a simple use-based explanation misses, and I will argue in this paper that this burden cannot be overcome. Given the breadth of this topic, what I present here will be a preliminary, if rather suggestive, investigation.

2 Theoretical Issues

In this section I discuss three characteristics of the CWC and investigate their implications for a synchronic-grammatical explanation. In the process a series of theoretical problems will be encountered which, I will argue, cannot be overcome.

2.1 The CWC crosses the syntax-morphology interface

A relationship between case and word order must be considered in terms of the syntax-morphology interface. If m-case affects word order within the synchronic grammar, then the syntax must depend (in part) on the morphology. This is not an uncommon assumption, implicit e.g. in the Lexicalist Hypothesis, according to which the morphological properties of a word, including its case-marking, are already determined at the point when it enters into the syntactic structure. Syntactic processes like those which determine word order could thus be sensitive to morphological factors. However, recent work, e.g. in Distributed Morphology, favors just the opposite view of the interface. There it is hypothesized that the syntax works with underspecified feature bundles, and morpho-phonological material is inserted into those bundles only after Spell-out. This is known as Late Insertion (see e.g. Halle and Marantz 1993; Marantz 1995). But if this is correct, then it is impossible for the presence or absence of case-marking to affect word order, because the markers themselves are not inserted into the structure until after word order has already been determined.

The obvious response to this objection is that what is relevant to word order is abstract Case, which is a syntactic feature and thus can have syntactic consequences, m-case just being the later spelling out of these features. But this too runs into a series of problems. The first is that the simplest hypothesis, that languages with m-case have syntactic Case, while those without it do not, is clearly untenable. Syntactic Case plays a role in DP positioning and licensing that is independent of word-order freedom and important in every language. Thus one would have to assume that abstract Case is universal, but fundamentally different in languages with m-case than in those without. But this would force Case theory to distinguish Latin Case from English Case while claiming that both are Case.

The second problem for a syntactic Case-based account is that there is mounting evidence **against** a direct connection between m-case and syntactic Case qua positional licensinge. The argument boils down to three points which I lay out here in extremely abbreviated form. For the full details I refer the reader to Yip et al. (1987), Marantz (2000) and Schütze (1997). a) The relationship between structural cases and structural positions is not one-to-one, but many-to-many. In most languages, the primary structural cases are assigned not to specific positions, but according to a sequence. Thus in a nominative-accusative language, the highest argument which does not receive a lexical case gets nominative, the next highest accusative. (See Yip et al. 1987 and Marantz 2000 for discussion and analyses of this phenomenon.) This accounts for why underlying objects are marked nominative in passives, unaccusatives and oblique subject constructions, whether they raise to subject position or not. b) It is possible for a position to be assigned m-case without being assigned syntactic Case, as in (5), where *einum* shows that the subject position gets quirky dative from *batna* in spite of being empty (Sigurðsson 1991):

(5) að PRO batna veikin einum er erfitt.
to PRO:DAT recover the-disease alone:DAT.MASC is difficult
`To recover from the disease alone is difficult.'

c) It is possible for an overt DP to be licensed in a position where m-case is not properly assigned. This is where default case shows up, like the nominative on the left-dislocated DP in (6) (see Schütze (2001) for discussion):

(6) Der/*Dem Hans, mit dem spreche ich nicht mehr.
the-NOM/*DAT Hans with him-DAT speak I not more
`Hans, I don't speak with him anymore.'

If case can be assigned by a default rule, then it can't license DPs, because then in any instance where normal case-assignment failed, the default rule would be able to apply, and no DP would ever go unlicensed, rendering the Case filter vacuous.

The third problem with a syntactic Case-based account is that it would not actually constitute a synchronic-grammatical explanation of the CWC (See Bobaljik to appear for similar discussion of theories of verb raising). It claims that word-order freedom depends directly on a special Case feature we can call K'. But of course K' is syntactic, so it cannot in turn depend synchronically on m-case (because in the view adopted here the syntax precedes the insertion of morpho-phonological material). We would have to say that m-case cues the acquisition of K', i.e. when the learner is presented with rich case-marking, she can conclude that the language must have K' and thus the relevant word-order freedoms. This itself is quite plausible, but it explains the CWC through principles of acquisition, **not** the synchronic grammar. An account of the CWC which is explicitly based on use and acquisition from the start does not run into any of these problems, because it does **not** rely on a synchronic causal connection from m-case to syntax.

2.2 The CWC involves optionality

The optionality that word-order freedom represents is notoriously problematic for certain approaches to the CWC. E.g. Roberts' (1997) theory boils down to the simple (and therefore attractive) idea that case-marking drives movement. That is, DPs move to Agr positions in order to unify in some sense with the case-marking which resides there, much as verbs have been argued to raise to positions in I to combine with the tense, mood and agreement affixes that reside there. However, this just predicts that languages with case-marking will have a different rigid word order than those without (e.g. IO-DO instead of DO-IO), while what we want to explain is that such languages allow multiple orders (e.g. IO-DO **and** DO-IO). Kiparsky (1997) and Neeleman and Weerman (1999), on the other hand, design their theories specifically to derive the optionality effect. For them, m-case does not cause DPs to move, but satisfies obligations which would otherwise have to be met by rigid ordering. Still, the theoretical steps they take to derive the optionality are themselves problematic. E.g. Neeleman and Weerman use the ECP to keep caseless DPs local to their licensing heads, formulating the ECP as a PF filter which rules out offending structures rather than as a principle which steers steps the derivation. However, both filters and the Government relation that the ECP depends on have been abandoned in recent work because of the complexity they introduce into the grammar. PF filters in particular are dubious because they are inconsistent with the default behavior characteristic of morphology. Similar remarks apply to the theory of Kiparsky (1997), although there matters are less clear because the framework he adopts is novel and less explicit on certain relevant details. Of course, optionality and how speakers deal with it is the proper domain of a theory of language use, so it should not be problematic for a use-based explanation of the CWC.

2.3 The CWC depends on 'richness and 'freedom'

Another familiar question for accounts of the CWC is precisely how 'rich' case-marking must be to allow word-order freedom, and how the grammar is to measure this 'richness'. How we answer this depends on where we encode the CWC. A synchronic-grammatical explanation must define it in purely formal terms that apply universally and cannot use speaker/hearer-based notions like 'sufficiently distinct', 'unambiguous' or 'salient', because such notions are not defined within the synchronic grammar. But a definition that can satisfy these restrictions **and** account for the cross-linguistic variation on this point turns out to be impossible to formulate.[2] First of all, richness cannot be reduced to the simple presence of a marker on a given form. It is typical in case-marking languages that one case in a given paradigm will have a zero marker, usually the nominative or absolutive, yet such forms have the same word-order freedom as ones with clear overt markers. What is actually relevant is distinctiveness between case forms, but even this cannot be determined by simply comparing the forms of a given noun. Latin *genu* 'knee' is the same in the nominative, dative, accusative and ablative, yet is no more restricted in its positional behavior than *princeps* 'chief', which has four distinct forms. So 'richness' of morphology must depend not on particular forms, but on distinctions between them, and is determined language by language, not word by word. But this is problematic if we want the synchronic grammar to do the determination, because it implies that it cannot be based strictly on the morphological information present in a given utterance. Rather, it must depend on paradigms having some reality in the synchronic grammar, which is commonly regarded as doubtful (Spencer1991, Bobaljik 2001).

Assuming for the sake of argument that this can be dealt with, we still need a universal formulation for the richness metric. We could propose something like the following: If a language distinguishes subject from object case in at least one number of at least one productive noun inflection class, then it has K'. This is intentionally analogous to a proposal made by Rohrbacher (1994): "A language has V to I raising if and only if in at least one number of one tense of the regular verbs, the person features [1st] and [2nd] are both distinctively marked (p.108)."

---

2 The following owes much to the discussion in Bobaljik (to appear) of the correlation between rich agreement morphology and verb raising. It is telling that both Bobaljik and Alexiadou and Fanselow (2000) come to the conclusion that the former cannot be explained within the synchronic grammar.

However, exceptions to Rohrbacher's principle are easy to find, including Tromsø Norwegian and Kronoby Swedish, which show no person distinctions, yet still have overt verb raising. Indeed, Bobaljik (to appear) has shown that no such definition could handle the variation attested within the Germanic languages, let alone universally. The task is no easier with the CWC, and in Section 3 we will see data that would contradict any definition of the CWC as a strong implication.

But even if we managed to construct a satisfactory definition, we would again have moved the CWC into the acquisition component, as was argued in Section 2.1. Consider what is at issue here. An essentially deductive step must be made from intra-paradigmatic morphological distinctions to the positing of a language-wide syntactic feature K'. Again, the synchronic grammar cannot make this leap, because it involves a syntactic feature depending on morphological information, contrary to the ordering of levels in DM. Rather, this is precisely the sort of thing the language acquisition device does, which constructs a grammar on the basis of a set of input data and the constraints of Universal Grammar. An explicitly use- and acquisition-based explanation of the CWC again avoids these problems, as it does not require a formal definition for 'rich' morphology. Rather, it depends on the ease with which speaker-hearers use and understand an utterance, thus it **can** use notions like 'sufficiently distinct' and 'unambiguous', which makes our task much easier.[3]

3 Empirical issues

There are thus serious theoretical issues for any explanation of the CWC within the narrow synchronic grammar. Still, the success of any theory hinges on its ability to account for the empirical facts, thus in this section I consider the data that have been used to argue for a synchronic-grammatical explanation.

3.1 How strong is the correlation?

It is clear that certain word-order freedoms do pattern with the presence of case-marking, but is the correlation a bi-conditional, a one-way implication, or just a tendency? There is no shortage of, e.g., languages like Icelandic and Grisons Swiss German which have extensive overt case-marking, yet lack scrambling. Kiparsky (1997) thus argues for a one-way implication:

[3] A related issue, which I cannot discuss here due to space constraints, is how a synchronic-grammatical account of the CWC can handle the clear effects of other factors, like agreement and intonation, on word-order freedoms. A use-based account correctly predicts without complication that all such factors which can aid interpretation will support marked word-orders, even in the absence of case-marking, while a synchronic-grammatical account must either formalize the affects of such factors in addition to those of case, or justify a more complicated hybrid account where case-marking constrains word-order freedom in the synchronic grammar but intonation does so in the use component. Similarly, a synchronic-grammatical account must give a principled explanation of why some reordering processes, like scrambling, seem to correlate with rich case-marking, while others, like topicalization, are more freely available.

> The most important point about this relationship is that it is...an exceptionless implication, which however holds in one direction only: *lack of inflectional morphology implies fixed order of direct nominal arguments* (p. 461, emphasis in original).

This is crucial, because as Kiparsky himself implies, only an actual implication can be properly encoded by a principle of the synchronic grammar. But even the one-way implication cannot be maintained. Dutch has no more case-marking than English, yet, as I will demonstrate in the next subsection, it allows scrambling of one object across another. Thus the CWC is not an implication, but a (strong) tendency. The synchronic-grammatical accounts cannot handle this. On the other hand, a tendential relationship is precisely what a use- and acquisition-based account predicts.

3.2 Dutch scrambling, Russian OVS, and pragmatic markedness

It is frequently reported that Dutch objects can scramble across negation and adverbs but not across other DPs. Kiparsky (1997), e.g., reports the following:

(7) dat Jan zijn vader het boek geeft
that Jan his father the book gives
'that Jan gives his father the book'

(8) *dat Jan het boek zijn vader geeft

However, things are not that simple. Zwart (1997) gives the data in (9)-(11) and the explanation below them:

(9) dat Jan Marie (gisteren) het boek gegeven heeft
that John Mary yesterday the book given has
'that John gave Mary the book yesterday.'

(10) ??dat Jan het boek Marie gegeven heeft
that John the book Mary given has
'that John gave Mary the book.'

(11) dat Jan het boek Marie terug gegeven heeft
that John the book Mary back given has
'that John gave the book back to Mary.'

> [10] is unacceptable in a neutral stress pattern, i.e. with *Marie* slightly focused. Almost any marked stress pattern makes [10] acceptable though. Thus, in [11] the particle *terug* is in focus, and the order of the objects appears to be free (p. 32).

In short, Dutch **does** allow scrambling of one object across another. A special intonation and an appropriate discourse context are required to license such orders, but this applies to German scrambling as well (see e.g. Lenerz 1977 and Haider 1993 for discussion). There is independent evidence that, where German has DO-IO order without special intonation, a different underlying syntactic structure is involved, which is analogous to the construction where the IO is marked in Dutch with *aan*, or in English with *to*. Kiparsky (1997) actually comes to this same conclusion and analyzes such examples differently from actual scrambling, following a view that is now fairly standard in analyses of the German double object (see e.g. Wegener 1991, Haider 1993, McFadden 2003).[4] In other words, the 'scrambling' without discourse and intonational motivation that is supposed to occur in German but not in Dutch is not scrambling at all, but alternation in base structures. The difference between German and Dutch is that German morphology spells out the post-accusative argument in the dative case while Dutch spells it out with a preposition. Of course, actual scrambling of DO over IO is much more common in German than in Dutch. This is for the very simple use-based reason that scrambling runs a greater risk of being misunderstood in the absence of case-marking. The lower frequency in Dutch implies higher pragmatic markedness, which we can plausibly interpret as there being a smaller number of contexts in which the scrambled order is felicitous. Dutch speakers presumably have difficulty when presented with this order in grammaticality tests because the correct context is difficult to supply out of the blue, and the marked intonation, unlike case-marking, is not indicated in writing.

Similar remarks are in order for claims that marked word orders are impossible even in languages with case-marking when the case-marking on the specific nouns involved is ambiguous. Jakobson (1936) e.g. claims that OVS order is possible in Russian only in sentences where the case-marking is unambiguous. Thus he gives (12) as OK, because *syna* is marked accusative, but says that (13) is impossible because neither *mat'* nor *doč'* shows a nominative/accusative distinction:

(12) Syna rodila mat' prošlym letom.
son:ACC bore mother:N/A last summer
'The mother bore the son last summer.'

(13) *Doč' rodila mat'.
daughter:N/A bore mother:N/A
intended: 'The mother bore the daughter.'

However, according to the native speakers I have consulted, this is actually incorrect for Russian, as are similar claims that have been made for German and other languages of the relevant type. What **is** true is that this order is marked and requires contextual motivation and/or a marked intonational pattern. Russian speakers thus presumably have the same difficulty with (13) encountered out of the blue that Dutch speakers have with (8).

[4] The idea is that certain verbs, including *verzeihen* 'forgive (s.o. s.t.)' have a basic structure in which the dative precedes the accusative. Others, including *aussetzen* 'expose (s.o. to s.t.)' have the accusative before the dative (see the cited works for syntactic and semantic tests which distinguish the two classes). Still others, including *geben* 'give' can appear in either frame. This is analogous to English, where verbs like *forgive* can only appear in the double object construction, others like *donate* can only appear in the *to*-dative, and still others like *give* can appear in either.

3.3 Misanalyses

To conclude this survey of empirical data, I will discuss an interesting class of examples which initially seem to support a synchronic-grammatical account of the CWC, but in the end turn out to have alternative analyses.

3.3.1 Ditransitives from Old to Middle English

OE lacked the *to*-dative construction (e.g. *John gave the books to Mary*), but allowed either both IO-DO and DO-IO orders without a preposition with the double object. In early ME, as case distinctions were being lost, the *to*-dative arose, and DO-IO order without a preposition disappeared. Thus it would seem that, when case-marking was lost, so was a certain type of word-order freedom, and the construction with *to* was created to fill the gap. However, I have argued elsewhere (McFadden, 2002), on the basis of evidence from a corpus study, that there were two distinct double object constructions in late OE and early ME. Sentences with IO-DO order had a structure like the modern double object construction, while at least some sentences with DO-IO order had a structure like the modern *to*-dative, with dative marking in place of the preposition *to,* just as is claimed above for German. In other words, DO-IO order was never lost. It simply became the *to*-dative construction via a morphological - not a syntactic - change.

3.3.2 'Semantic' case DPs

DPs bearing 'semantic' or 'adverbial' case show a remarkable degree of freedom in their positioning cross-linguistically, frequently beyond that of other case-marked DPs in a given language. In German, e.g., argument DPs are banned from the position following the participle, as in (15), but semantic case-marked DPs, like the accusative of time in (16) are OK here, if somewhat marked:

(14) Er hat letzten Sonntag seinen Bruder besucht.
he has last Sunday:ACC his brother:ACC visited
'He visited his brother last Sunday.'
(15) *Er hat letzten Sonntag besucht seinen Bruder.
(16) ?Er hat seinen Bruder besucht letzten Sonntag.

It is tempting to think that it is the special case-marking that frees these nominals from normal constraints on DP ordering. But Emonds (1987) and Nikanne (1993) have argued that these are underlyingly PPs realized morphologically as case-marked NPs. This accounts for how they are assigned θ-roles, why they have the same semantic range as overt PPs, a number of restrictions on their behavior with respect to binding and secondary predication and the word-order freedom just noted. As (17) shows, the ability to extrapose is a general property of German PPs:

(17) ?Er hat seinen Bruder besucht im März.
he has his brother visited in March

'He visited his brother in March.'

The word order freedom is thus caused by the fact that these constituents are PPs. The case-marking is a symptom of the PP structure, not the cause of the freedom.

3.3.3 Case drop

The clearest possible evidence for a synchronic CWC would be if some word order were only possible in a language in the presence of a case-marker. Examples of this kind have been reported by Lamontagne Travis and (1987) for Japanese and Turkish (their behavior is the same, so I will restrict attention to the former):

(18) John-ga dare-(o) nagutta no?
John-NOM who-(ACC) hit Q
'Who did John hit?'

(19) Dare-*(o) John-ga nagutta no?

In the unmarked order in (18), the accusative marker *-o* on the direct object can be dropped, but if the object is fronted, as in (19), it cannot. Lamontagne and Travis thus argue that case is a syntactic head, subject to the ECP when it is phonologically null. The null K is properly governed in (18), but not in (19), because it has moved away from the licensing verb.

However, an alternative analysis is available which is fully consistent with my hypothesis and avoids the theoretical problems of Section 2. Recall that there is nothing in my theory to keep word order from affecting m-case. Like all morphology, it is inserted on the basis of the output of the syntactic derivation. Thus all we need say about (18) and (19) is that in Japanese (and Turkish), a case-marker **must** be inserted on objects that have moved out of VP (or whatever the precise generalization turns out to be). When the object remains VP-internal, there are two morphological options, *-o* and *-Ø*, with the choice between the two being conditioned by other factors, apparently related to focus. In fact, there are reasons to prefer the latter account. In particular, the former would predict object topicalization to be impossible in languages which lack m-case, because there would be no overt version of K available to obviate the effects of the ECP. But of course object topicalization is very much possible in in many languages without m-case, like English and Dutch. Of course, if case-drop depends on word-order, then topicalization in caseless languages presents no problem.

4 Conclusion

So we have seen theoretical and empirical arguments that the CWC is a tendency that results from principles of language use, acquisition and change, not a principle of the narrow synchronic grammar. We can imagine that the correct explanation will be roughly as follows. As a language like Old English with a high degree of word-order freedom lost its case-marking, the more marked word orders would have become increasingly difficult to interpret correctly. As a result, it would have become increasingly difficult for children to acquire the processes that derived such marked word orders. This would have been helped along by increasing avoidance of the use of marked orders by those who had acquired them, to ensure being understood. The marked orders would have become even more marked, requiring stronger and stronger pragmatic and intonational motivation to ensure their unambiguous interpretation. Eventually, the evidence for the marked orders in the primary linguistic data would have been insufficient to allow children to acquire the word-order freedom, yielding essentially the Modern English situation. Of course, other factors like agreement and intonation can play a significant role in the interpretation of word-order variation, allowing in certain cases the preservation of word-order freedom even in the absence of morphological case-marking, as in Dutch.

References

Alexiadou, Artemis and Gisbert Fanselow. 2000. On the correlation between morphology and syntax: the case of V-to-I. In *Proceedings of GLOW* 2000.

Bobaljik, Jonathan. 2001. Syncretism without paradigms. In Geert Booij and Jaap van Marle (eds.). *Yearbook of Morphology 2001*. Dordrecht: Kluwer.

Bobaljik, Jonathan. to appear. Realizing Germanic inflection: why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics.*

Emonds, Joseph. 1987. The invisible category principle. *Linguistic Inquiry* 18:613-632.

Haider, Hubert. 1993. *Deutsche Syntax - generativ*. Tübingen: Gunter Narr Verlag.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds.). *The view from building 20.* Cambridge, Mass.: MIT Press.

Jakobson, Roman. 1936. Beitrag zur allgemeinen Kasuslehre. *Travaux du Cercle linguistique de Prague* 6.

Kiparsky, Paul. 1997. The rise of positional licensing. In A. van Kemenade and N. Vincent (eds.) *Parameters of morphosyntactic change*. 460-494. CUP.

Lamontagne, Greg and Lisa Travis. 1987. The syntax of adjacency. In *Proceedings of WCCFL* 6:173-186.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen.* Tübingen: Gunter Narr Verlag.

Marantz, Alec. 1995. A late note on late insertion. In Young-Sun Kim (ed.) *Explorations in generative grammar,* 396-413. Seoul: Hankuk Publishing Co.

Marantz, Alec. 2000. Case and licensing. In Eric Reuland (ed.) *Arguments and case: explaining Burzio's Generalization*. Philadelphia: John Benjamins. Originally appeared in *Proceedings of ESCOL* 1991.

McFadden, Thomas. 2002. The rise of the *to*-dative in Middle English. In David Lightfoot (ed.) *Syntactic effects of morphological change*. Oxford: OUP.

McFadden, Thomas. 2003. The underlying structures of German inherent Case. Presented at the joint meeting of the FGLS and the SGL.

Neeleman, Ad and Fred Weerman. 1999. *Flexible syntax: A theory of case and arguments*. Dordrecht: Kluwer.

Nikanne, Urpo. 1993. On assigning semantic cases in Finnish. In A. Holmberg and U. Nikanne (eds.) *Case and other functional categories in Finnish syntax*. New York: Mouton de Gruyter.

Roberts, Ian. 1997. Directionality and word order change in the history of English. In A. van Kemenade and N. Vincent (eds.) *Parameters of morphosyntactic change*, 397-426. Cambridge University Press.

Rohrbacher, Bernhard Wolfgang. 1994. The Germanic VO languages and the full paradigm: a theory of V to I raising. Doctoral Dissertation, UMass.

Schütze, Carson. 1997. Infl in child and adult language: Agreement, case and licensing. Doctoral Dissertation, MIT.

Schütze, Carson. 2001. On the nature of default case. *Syntax* 4:205-238.

Sigurðsson, Halldór Ármann. 1991. Icelanidic Case-marked PRO and the licensing of lexical arguments. *Natural Language and Linguistic Theory* 9.

Spencer, Andrew. 1991. *Morphological theory*. Oxford: Blackwell.

Wegener, Heide. 1991. Der Dativ - ein struktureller Kasus? In G. Fanselow and S. Felix (eds.) *Strukturen und Merkmale syntaktischer Kategorien*, 70-103. Tübingen: Narr.

Yip, Moira, Joan Maling and Ray Jackendoff. 1987. Case in tiers. *Language* 63:217-250.

Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrect: Kluwer.

Thomas McFadden
619 Williams Hall
Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104-6305

tmcfadde@ling.upenn.edu

Subsumption in Feature Theory and Speech Recognition

MORITZ NEUGEBAUER
University College Dublin

0. Introduction

Most phonological approaches to the internal structure of speech sounds agree on the general assumption that the layer known as the *segmental layer* is to be interpreted as an abbreviated notational variant of a fully specified feature set, and that this set of features has a complex inherent internal structure. Various graphical formalisms have been applied to visualize dependencies among phonological features ranging from feature tables over tree structures to attribute-value matrices. However, in this paper it will be shown that neither pair of these representations is equivalent from a formal point of view. This paper claims that feature tables as well as feature trees lack a formal definition of *subsumption* establishing a feature hierarchy apart from the graphical notation itself. Therefore, a novel feature theory is proposed, motivating the notion of a subsumption relationship among supersets and subsets of features by (partial) class behaviour. We define how to represent features "if they consistently behave as a unit" (Clements 1985:225).

The paper is organized as follows: section 1 introduces representations which are common in linguistic approaches to structure below the segmental level while simultaneously providing for a critical review regarding the interpretation of the representational means in use. By departing from work on exclusively phonological feature structure, the principal interest of this paper turns to a notion of *feature descriptions* which is intended to establish a logic which allows statements on the decidability and consistency of actual feature terms to be made (cf. section 2). On this basis, section 3 sketches applications of the previously developed feature theory in actual speech recognition systems, focussing on issues of the representation of featural dependencies in the lexicon. The paper concludes with a brief summary and a discussion of topics for future research.

1. Subsegmental Phonological Representations and their Interpretation

An extensive body of research, largely independent from constraint-based or rule-based assumptions, builds on a representational layer below the segment. Recently, the observation that the set of phonological features is not free of but rather determined by featural interdependencies has attracted remarkable attention. In the following, the representational apparatus of previous theoretical approaches will be revised and it will be subject to significant modifications to capture basic logical relations between the smallest units of phonological description. Beginning with the concept of binary feature specification visualized in feature tables, the forthcoming subsections explore the expressivity of phonological feature representations from a logical point of view. The review proceeds to the characterization of feature trees and finally of attribute-value matrices as applied in numerous prior analyses couched in computational phonological frameworks.

1.1 Phonological Tables

According to assumptions initially formulated in Linear Phonology (cf. Chomsky/ Halle 1968), phonological representations consisted of a concatenated bundle of (mostly) articulatory features which are specified in a binary fashion. The internal organisation, i.e. the data structure of these feature bundles in this conception should not to be mistaken for a linked list as it lacks for example the ability to reflect even local precedence relations between individual features. In particular, the linear ordering in visual representations is completely arbitrary and therefore not applicable as a structural reference in the formulation of phonological rules. On the other hand, without resorting to evidence from non-phonological information in a strict sense such as articulatory phonetics, the identical behaviour of two features captured in a phonological rule always seems to occur accidentally. The feature table below gives a feature table representation for a three-member set of commonly used features:

(1)

voiced	±
spread	±
constricted	±

A model of feature organisation of this particular kind makes false predictions whilst claiming that features can be combined without any constraints reflecting the set of possible feature combinations (i.e. sounds) in natural languages and/or constraints establishing classes of features e.g. for access by the rule component (cf. Broe 1992:155). Assuming that the feature set given in (1) is subject to assimilation processes in a certain language, the above representation neither suggests similar phonological behaviour of the features in question, nor does it seem possible to associate just these features – apart from the rest of a remaining feature bundle – with two adjacent skeletal positions. The natural class of laryngeal features, which is generally assumed to consist of exactly the feature set depicted in (1), can only be captured by the invention of an additional feature [±laryngeal] along with a set of implication rules of the kind: [+voiced] → [+laryngeal]. However, this solution does not provide us with an elegant foundation for phonological classification since feature classes can only be described by simultaneously inventing new features (such as [±laryngeal]). In addition, these features would lack any segmental equivalent in a phonological system, given that there is no such sound whose place of articulation is sufficiently described by [±laryngeal] since in any case further rules would have to be applied. The subsumption relation encoded can be paraphrased as "every [+voiced] segment is also [+laryngeal] but not vice versa".

Leaving feature table representations behind, the following section focusses on the contributions of Feature Geometry to the task of phonological classification also referred to as non-linear representations.

1.2 Phonological Trees

The previous subsection drew attention to the observation that feature tables cannot encode local precedence relations as in linked lists for example. However, following the lines of research in Autosegmental Phonology (Archangeli 1985) exactly these relations were captured within a coplanar model of featural organization as given in (2). The representation can be interpreted as a non-branching tree stating that except for feature A, every feature depends on the presence of further features.

(2)

```
     C
     |
 Feature 1
     |
 Feature 2
     |
 Feature 3
```

The strong claims made in this model appear to be too restrictive for the phonological domain: there is no such phonological feature which is obligatory for every C-element in a given language. The related multiplanar approach, which is based on the idea that any individual feature can behave independently of the rest of the feature set on an individual representational layer faces the same problems as the linear model since no means for featural classification are provided. Table (3) gives a sample representation:

(3)

```
           Feature 1
               |
 Feature 2 _   C   _ Feature 4
             – | –
           Feature 3
```

As a model which combines both feature dependencies and independencies, the inspiration for Feature Geometry dates back to *The Sound Pattern of English*

(Chomsky and Halle 1968) where the heterogeneous nature of unordered feature bundles was already noticed. Based on the observation that "it is less apparent but nevertheless true that the feature matrix formalism incorporates certain implications for feature organization that do not follow from the vaguer notion of bundle" Clements (1985:225) elaborates the model of Feature Geometry which introduces class nodes in addition to phonological features. These nodes define feature classes which hence can take part in phonological processes in a homogeneous way. Motivated by the work on the hierarchical organisation of phonological features by Clements (1985), Lahiri and Reetz (2002) apply a tree-based organisation for place of articulation features in their FUL (fully underspecified lexicon) system for speech recognition. The important question with regard to the topic of this paper is in how far class nodes such as *Articulator* and *Aperture* improve a feature system like the one reviewed in the previous subsection where feature classes are captured by the means of additional features.

According to Clements (1985:229) there is a clear-cut distinction between class nodes and features. Whereas the concept of phonological features remains similar to the SPE in the sense that features specify segmental properties, the content of class nodes differs remarkably since "their feature content is entirely determined by the features which they dominate". This states that class nodes cannot exist on their own but that they rather define class membership relations among sets of phonological features (cf. Bird 1991:137).

A very important aspect that follows from the above definition is that a class node is in contrast to a feature not specifiable but can rather be defined as the set of the contexts of its daughters (cf. Kornai 1994:26). This property can be derived and does not need to be motivated on the basis of empirical investigations, by an explicit activation condition for class nodes (cf. Avery and Rice 1989:183), by some theory-specific reformulation (cf. Padgett 2002) or by the assumption that class nodes are privative features (cf. Wiese 2002). It can be added, that in the latter argumentation the same strategy as in linear phonology is applied since new features are proposed to classify (at least) two others.

Another common graphical notation for feature hierarchies besides features trees is the attribute-value matrix which is described in more detail in the following section. Such matrices provide a visualization for *feature descriptions* as they are applied in the forthcoming core part of the paper. Once the feature formalism is developed it will be applied in finite state phonology for the purpose of speech recognition (cf. section 3).

2. Feature Descriptions

In constraint-based grammar formalisms and in logic programming languages various feature descriptions are employed and feature descriptions date back to founding work on Functional Grammar (cf. Kay 1979), though for the rest of this paper we will follow the formalizations in Backofen and Smolka (1995).

The common notational primitives of feature descriptions are functional attributes called features. This notion of *feature* crucially differs from its use in the previous sections since from now on we will distinguish between *sorts* and *features* while phonological features as used before are sought to correspond most closely to sorts.

The descriptions considered in this paper are obtained from a signature of binary and unary predicates called features and sorts, respectively. Built over the sort alphabet and the feature alphabet, feature descriptions are first-order formulae where in admissible interpretations features must be functional relations and distinct sorts must be disjoint sets. A feature description written in attribute-value matrix (AVM-) notation is the one in (4) a. It is stated that there exists a segment x which carries a place feature, more precisely a height feature LOW whose value is specified as negative. While PLACE and LOW are sorts, *segment* and *height* are encoded as features. It should be noted both that sorts can take complex values (cf. PLACE) and that each bracket in an AVM is indexed with a feature. That index would e.g. define that LOW can take boolean values only. Of course, each matrix can be written in plain first-order syntax as shown in (4) b.

(4) a.

$$ {}_{segment}\left[\text{PLACE} \; {}_{height}\left[\text{LOW} : \; - \right] \right] $$

b.

$$\exists\, P \,(\, \text{segment}(x) \wedge \text{place}(x,P) \wedge \text{height}(P) \wedge \text{low}(x,-)\,)$$

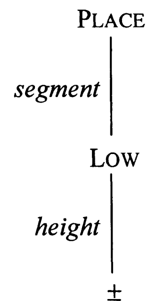
This paper does not present the first attempt to capture phonological feature structures in a strict sorted feature logic. For instance, Bird and Klein (1994) developed a prosodic type hierarchy along similar lines. However, one striking difference between the two approaches seems to be the missing feature indication in the AVM below, where none of the internal brackets comes with further featural information (cf. Neugebauer to appear).

(5)

$$ {}_{segment}\left[\begin{array}{ll} \text{LARYNGEAL :} & \left[\begin{array}{ll} \text{SPREAD :} & \text{boolean} \\ \text{CONSTRICTED :} & \text{boolean} \\ \text{VOICED :} & \text{boolean} \end{array} \right] \\ \text{SUPRALARYNGEAL:} & \left[\begin{array}{ll} \text{MANNER :} & \left[\begin{array}{ll} \text{NASAL :} & \text{boolean} \\ \text{CONTINUANT :} & \text{boolean} \\ \text{STRIDENT :} & \text{boolean} \end{array} \right] \\ \text{PLACE :} & \left[\begin{array}{ll} \text{CORONAL :} & \text{boolean} \\ \text{ANTERIOR :} & \text{boolean} \\ \text{DISTRIBUTED :} & \text{boolean} \end{array} \right] \end{array} \right] \end{array} \right] $$

A notational variant of feature descriptions are feature graphs. Edges in feature graphs are labelled with features while nodes are labelled with sorts and thus feature descriptions are interpreted over feature trees in the following way: every sort symbol is taken as a unary predicate where a sort constraint $A(x)$ holds if and only if the root of the graph x is labelled with A. Contrastively, features symbols are taken as binary predicates where a feature constraint $f(x,y)$ holds if and only if the graph x has the direct subgraph y at feature f. While we can represent the first-order term in (4) b. as a feature graph in a straightforward fashion (cf. (7)), it should be obvious that missing feature labels as in (5) would lead to an ill-formed feature graph which would have to appear on the arcs of a well-formed graph structure. We already stated that the edges of the graph are called features whilst this is of course only observable in complex graphs of feature-value pairs where each value is either again a complex graph or an atomic one. An atomic graph is just a symbol; that is, it contains no features. In figure (6) the leaf of the graph is an atomic value, i.e. "-". It is this basic aspect of feature graphs which disallows the appearance of specified class nodes as mentioned in section 1.2. where complex graphs such as PLACE in (6) are mistreated as atomic graphs.

(6)



In these preliminary remarks on feature graphs the notion of *subgraph* has already been introduced. This aspect as well as the basic axioms of our feature theory are to be defined in the subsequent sections. Since all the following builds on predicate logic, the necessary notions are recalled briefly to conclude this introductory section.

We assume a set SOR of unary predicate symbols called sorts and a set FEA of binary predicate symbols called features. Given a feature structure, a sequence of labels is used to extract a substructure. Such a sequence of features is called a path over the set of all features. The symbol ε denotes the empty path which means that for every path p there exists a set of equivalent paths $\varepsilon p = p = p\varepsilon$. In the first case, ε is a prefix of p meaning that there exists a path p' such that $\varepsilon p' = p$. In addition to the partially disjoint sets SOR and FEA we assume an equally disjoint set of variables. Under our signature SOR $\cup$ FEA, every term is a variable and an atomic formula is either a feature constraint $f(x,y)$, a sort constraint $A(x)$, an equation $x = y$, $\top$ or $\perp$. In all these cases x, y denote variables, A denotes a sort and f a feature. We speak of a theory as a set of closed formulae where a model of a theory is a structure that satisfies every formula of a theory. The following section defines a theory by means of three axiom schemes. These axiom schemes are inspired by work the of Backofen and Smolka (1995).

2.1. Axioms

As stated at the beginning of section 2, it is assumed that features are functional which can be demonstrated given two instances *f(x,y)* and *f(x,z)* of a feature constraint *f*, where *x*, *y*, *z* are variables. If both constraints are supposed to be elements of an identical formula ϕ then *y* and *z* must be identical:

(7) $$\forall x \forall y \forall z(f(x,y) \wedge f(x,z) \rightarrow y = z) \text{ for every feature f}$$

Sorts are required to be mutually disjoint sets. Thus the second axiom scheme says that if a sort constraint *A* holds for a variable x in a formula ϕ and there is a sort constraint *B* over *x* defined in the same formula, then *A* and *B* must be identical. We capture this by stating the second axiom as in (8) which defines that the statement that a variable *x* is as well of sort *A* and *B* is false.

(8) $$\forall x(A(x) \wedge B(x) \rightarrow \perp)$$

Thus far both axiom schemes are formulated on the general assumption that all features and sorts are non-empty and that feature descriptions are consistent and have solutions. To actually guarantee these properties of our feature theory a third axiom scheme is required stating that consistent feature descriptions are satisfiable. For this reason we introduce the notion of a solved clause to describe the satisfiability of feature descriptions. A solved clause can be seen as the graph whose nodes are the variables appearing in the clause and whose arcs are given by the feature constraints. Nodes are labelled by sort constraints or by exclusion constraints which state a feature *f* is undefined on a variable *x*. We will represent exclusion constraints as $xf\uparrow$ saying that *f is undefined on x*. Formally a solved clause is a conjunction σ of atomic formulae i.e. feature and sort constraints as well as exclusion constraints. The latter kind, the exclusion constraints, are equivalent to $\neg\exists y\ f(x,y)$ for some variable $y \neq x$. A variable *x* is constrained in a solved clause σ if σ contains one of these constraints on *x*. Based on these notions we can finally state the third axiom scheme: $\exists X\sigma$ (for every solved clause σ and $X = CV(\sigma)$).

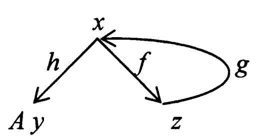
The feature theory developed thus far is the set of all sentences that can be obtained as instances of the axiom schemes in (8 – 10). According to these three axiom schemes, a solved clause σ must satisfy the following four conditions. The most obvious condition is that no atomic formula such as an equation, sort or feature or an exclusion constraint is supposed to occur twice in σ. In addition, sorts must be mutually disjoint sets, features are supposed to be functional and finally if a feature is defined on a specific variable then it cannot be the case that there is an exclusion constraint on this specific variable and this specific feature in an identical solved clause σ. An example of a graph representation of a solved clause will be given in the following section since it focusses on models of our feature theory in terms of feature graphs.

2.2. Feature Graphs

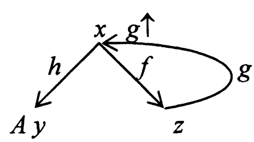
Building on the axioms defined above, it is the aim of this section to provide for a model of our feature theory which allows us to represent feature structures – here – in the area of phonology. The graph representations which show a certain degree of similarity to the tree diagrams commonly used in Feature Geometry. Thus, we introduce a representational formalism which provides us with illustrative means for application in section 3.

As we introduced our three pairwise disjoint sets of variables, sorts and features it was noted that we can extract a substructure of a feature description using a sequence of labels. Such a sequence of labels is called *path* and is defined as an expression over FEA*, the set of features. Before we proceed to give a graph representation of a solved clause, we introduce the notion of *feature pregraph* which is a pair of a solved clause σ as defined in the previous section and a variable x. In the case of feature pregraphs, the solved clause does not contain exclusion constraints such that for every variable y occurring in σ, there exists a path p from x to y in the solved clause σ (i.e. $p(x,y) \in \sigma$). A feature pregraph is called *subpregraph* of a feature pregraph if the solved clause of the pregraph subsumes the solved clause of the subpregraph and the two variables in question are identical or x is an element of the set of variables defined in the subpregraph. Note that feature graphs are directed graphs with a distinguished root node.

(9) a.



b.



c. $h(x,y) \wedge f(x,z) \wedge g(z,x) \wedge gx\uparrow \wedge A(y)$

In the example (9) a. the variable y carries the sort symbol A whereas the nodes labelled with x and z are not specified in a similar way. This is possible only in feature graphs whilst in feature trees each node must carry a sort symbol (cf. Backofen and Smolka 1995). In addition to the pregraph in (9) a., a feature graph contains the exclusion constraints as mentioned above. The above table contains the graph representation in (9) b. as well as the corresponding solved clause in (9) c.

3. Applications in Speech Recognition

This last section argues that subsumption relations between finite state automata modelling phonotactic regularities motivate a set of feature graphs for lexical knowledge representation in terms of supergraphs and subgraphs, respectively. Since we aim to define transitions over sets of sounds rather than atoms we make reference to a specific line of research in the field of finite state methods in natural language processing, known as *predicate augmented finite state automata* (in the following *pfsa*; cf. van Noord/Gerdemann 2001). This will be sketched in section 3.1. Our novel approach to phonological feature descriptions in combination with the previously mentioned augmented finite state techniques will build on the model of *Time Map Phonology* (cf. Carson-Berndsen 1998), a constraint-based model of speech recognition. Since this model implements restrictions on syllable well-formedness in terms of finite state techniques by defining phonotactic automata, we present a homogenous model in section 3. We show that we can finally derive complex feature graphs from a regular grammar for a selective data set of laryngeal features in German. Consequently, language-dependent phonological feature classes are presented as a result. These feature classes are regarded as a significant contribution to lexical knowledge in computational recognition systems since they allow to further determine the phonotactic context. We agree with Fant (1970:210) that feature classes constitute an essential component of the phonological lexicon.

3.1. Integrating Feature Descriptions and Finite state models

Most recently, some approaches have proposed finite state automata where transitions are associated with sets of symbols (cf. Bird and Ellison 1994, Walther 1999). Partially motivated by these developments van Noord/Gerdemann (2001) present finite state machines in which atomic symbols on the transition arcs are replaced by arbitrary unary predicates. In our feature theory unary predicates are used for a very similar purpose, namely as sort constraints labelling the nodes of feature graphs. Below we briefly describe the concept of predicate augmented finite state automata (called *recognizers* by van Noord/Gerdemann 2001) and consequently replace their set of arbitrary predicates with a set of sort constraints.

In contrast to a standard finite state automaton, a *pfsa A* is specified by a six-tuple $(\Delta,\Sigma,\Pi,E,S,F)$ just as described below:

(10)

- the finite set of states Δ
- the input alphabet Σ
- Π the set of predicates over Σ
- E a finite set of transitions $\Delta \times (\Pi \cup \{E\}) \times \Delta$
- the set of start states $S \mid S \subseteq \Delta$ and
- the set of final states $F \mid F \subseteq \Delta$

We can extend E to a function E^ and define the relation $E\hat{} \subseteq \Delta \times \Sigma^* \times Q$ by induction. For this purpose we state in a first step that for all states q in Δ there exists an ε-transition which results in the same state q. Similarly, for all transitions (p,ε,q) in E^ there exists a corresponding transition (p,ε,q) in E. A pfsa is called ε-free if there are no $(p,\varepsilon,q) \in E$. For all other transitions (q_0,π,q) in E and for all $\sigma \in \Sigma$ if $\sigma(\pi)$ there exists a corresponding transition (q_0,π,q) in E^. As a last step, if (q_0,x_1,q_1) and (q_1,x_1,q_1) are both in E^ then we can infer that $(q_0, x_1 x_2, q_1)$ is also an element of E^.

Finally, the language accepted by A is defined to be $\{\omega \in \Sigma^* \mid q_S \in S, q_F \in F, (q_S, \omega, q_F) \in E^\wedge\}$ (cf. van Noord/Gerdemann 2001:266).

Our extension now consists in replacing Π, the set of predicate symbols with a set of sort constraints which comes equipped with all properties defined in section 2.1. One significant property concerns the determinization of automata where van Noord/Gerdemann mention on the one hand well-known aspects like the requirement of just single start state and like the existence of at most one transition (q,π,q') for all states and symbols such that $\sigma(\pi)$. On the other hand they have to exclude overlapping transitions, i.e. intersecting sets of segments. To exclude these cases where more than one transition is applicable one has to compute appropriate individual transitions for all subsets. Thus for two such individual overlapping transitions over sets A and B in a *pfsa* we finally end up with three transitions for $A \wedge B$, $\neg A \wedge B$ and $A \wedge \neg B$. In our approach, this computation can hugely benefit form our inherently structured graph representations and the related possibility to extract subgraphs. Simply by reference to (possibly negated) sets we provide for the predicates which serve as labels for each transition in the automaton. Inspired by van Noord/ Gerdemann, we compute the transitions over a set SOR of sort constraints leaving a given subset Δ. In a first step, we compute the function $Trans^{\Delta} : SOR \rightarrow 2^{Q}$, defined as $Trans^{\Delta}(\pi) = \{q \in Q \mid p \in \Delta, (p, \pi, q)\} \in$ E. Thus for instance, suppose $\Delta = \{q\}$ along with the transitions

(11) $$E = \{ (p,\pi_1,q_1),(p,\pi_1,q_2), (p,\pi_2,q_2), (p,\pi_2,q_3), (p,\pi_2,q_4) \}$$

Consequently, the logical combinations of sets of states for each of the sort constraints π_1, π_2 are the following:

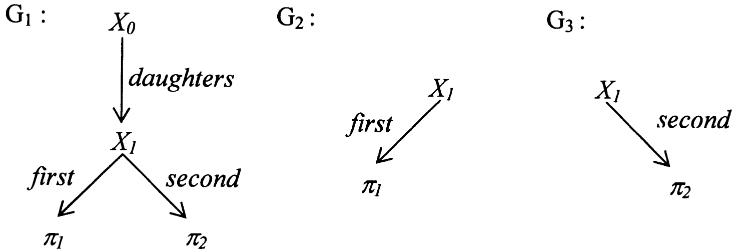
(12) $$\mathrm{Trans}^{\Delta}(\pi_1) = \{q_1,q_2\},\ \mathrm{Trans}^{\Delta}(\pi_2) = \{q_2,q_3,q_4\}$$

We assume that SOR' is the set of sort constraints in the domain of $Trans^{\Delta}$ and receive the following transitions over sorts as a result. Since we are dealing with two sort constraints as predicates and a negation operator, four possible combinations have to be accounted for.

(13) $$\begin{gathered} (\Delta, \pi_1 \wedge \pi_2, \{ q_1,q_2,q_3,q_4 \}) \\ (\Delta, \neg\pi_1 \wedge \pi_2, \{ q_2,q_3,q_4 \}) \\ (\Delta, \pi_1 \wedge \neg\pi_2, \{ q_1,q_2\}) \\ (\Delta, \neg\pi_1 \wedge \neg\pi_2, \{\varnothing\}) \end{gathered}$$

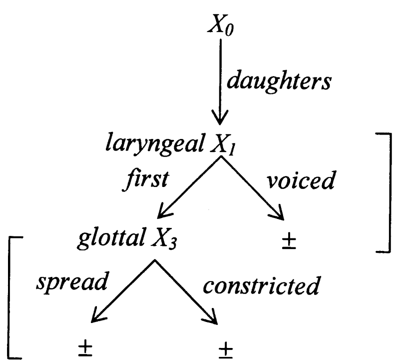
The contribution of feature descriptions and corresponding feature graphs can be demonstrated by the three related graphs in (14). Graph G_1 is the most specific one and therefore represents the simple conjunction of the sort constraints π_1, π_2. G_1 subsumes G_2 and G_3 while the latter two graphs are not connected in any subsumption relation. The negation of both constraints naturally does not meet any graph representation, at least not for the given predicates π_1, π_2.

(14)



To exemplify the predicate augmented finite state approach as described above we give the following complex laryngeal tree including the three phonological features introduced in (1).

(15)



For a precise application of this feature graph structure for determinization purposes of finite state machines we refer to a computation along the lines of (11 - 13).

4. Summary and Concluding Remarks

Based on a review of representational means commonly used in phonological feature theory, this paper presented a novel approach in terms feature descriptions for the internal structure of speech sounds. For this purpose we concentrated on feature theory in constraint-based grammar formalisms and in logic programming languages to define a theory of feature descriptions and a corresponding model of feature graphs.

Equipped with these foundations we integrated our representations with finite state techniques and concluded with an exemplification in the field of speech technology.

Since the paper primarily focussed on logical refinements rather than the analysis of speech data the latter has to be a major topic in future research. Another desirable issue is to explore our approach for finite state transducers rather than automata since various applications in speech recognition rely on these machines (cf. Carson-Berndsen 1998). Special attention has to be given to determinization and minimization of transducers as pointed out in van Noord/Gerdemann (2001).

5. Acknowledgements

This material is based upon works supported by the Science Foundation Ireland under Grant No. 02/IN1/I100. The opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Science Foundation Ireland.

References

Archangeli, D. 1985. Yokuts harmony: Evidence for coplanar representation in nonlinear phonology. *Linguistic Inquiry* 16: 335-372.

Avery, P. & K. Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6: 179-200.

Backofen, R. & G. Smolka. 1995. A complete and recursive Feature Theory. *Theoretical Computer Science* 146: 243-268.

Bird, S. & E. Klein. 1994. Phonological Analysis in Typed Feature Systems. *Computational Linguistics* 20: 455-91.

Bird, S. & M. E. Ellison. 1994. One-Level Phonology: autosegmental representations and rules as finite automata. *Computational Linguistics* 20: 55-90.

Bird, S. 1991. Feature structures and indices. *Phonology* 8: 137-144.

Broe, M. 1992. An introduction to feature geometry. In G. J. Docherty & D. Ladd (eds.) *Papers in Laboratory Phonology*,149-165. Cambridge: Cambridge University Press.

Carson-Berndsen, J. 1998. *Time Map Phonology: Finite State Models and Event Logics in Speech Recognition.* Dordrecht, Holland: Kluwer.

Chomsky, N. & M. Halle. 1968. *The Sound Pattern of English.* New York: Harper and Row.

Clements, G. N. 1985. The Geometry of Phonological Features. *Phonology Yearbook* 2: 225-252.

Fant, G. 1970. Automatic Recognition and Speech Research. Reprinted in G. Fant 1973. *Speech Sounds and Features*, 210-215. Cambridge, MA: MIT Press.

Kay, M. 1979. Functional Grammar. In *Proceedings of Fifth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA : UC Berkeley.

Kornai, A. 1994. *Formal Phonology*. Garland Publishing.

Lahiri, A. & H. Reetz. 2002. Underspecified recognition. In C. Gussenhoven & N. Werner (eds.), *Laboratory Phonology VII*, 637-675. Berlin: Mouton.

Neugebauer, M. to appear. Computational Phonology and Typed Feature Structures. In *Proceedings of CamLing*. Cambridge: University of Cambridge.

Padgett, J. 2002. The unabriged feature classes in Phonology. *Language* 78.1: 81-110.

van Noord, G. & D. Gerdemann. 2001. Finite State Transducers with Predicates and Identities. *Grammars* 4 (3): 263-286.

Walther, M. 1999. One-Level Prosodic Morphology. *Marburger Arbeiten zur Linguistik* 1. University of Marburg.

Wiese, R. 2002. *The Phonology of German*. Oxford: Oxford University Press.

Moritz Neugebauer
Department of Computer Science
University College Dublin
Belfield, Dublin 4
Republic of Ireland

moritz.neugebauer@ucd.ie

A bipartite verb stem outlier in Eurasia: Nakh-Daghestanian

JOHANNA NICHOLS
University of California, Berkeley

1. Bipartite stems

A bipartite verb stem (Jacobsen 1980a, DeLancey 1996) is one that is segmentable into two parts on certain morphological criteria but behaves in most respects like a simple root. Bipartite stems are not the same thing as compounds, though in most languages with bipartite stems some of the more transparent bipartites are also amenable to analysis as compounds. The definition of *bipartite stem* used here is: a segmentable simplex stem; or a stem with inflection positioned so as to split the stem into two parts.

The first element of a bipartite stem is often called a preverb. Here, however, I will simply speak of the *initial* element and the *final* element of a bipartite stem, so as not to impose any assumptions about which piece is more predicate-like and which is more likely to be an open-class element. In fact, cross-linguistically, both the initial element and the final element are prone to be closed classes, and both are likely to include many dependent stems and some cranberry morphs.

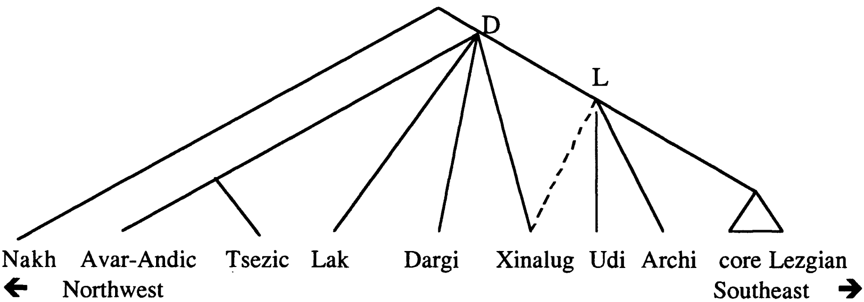
I will use the term *interposition* to describe the positioning of inflectional material between the two pieces of a bipartite stem. (This is different from infixation, which occurs inside of a simple stem, where its position is usually defined phonologically.) In some languages bipartite stems appear to have arisen when clause elements became incorporated, trapping inflectional morphemes that had formerly been ordinary prefixes or suffixes on the incorporating element. The purpose of this paper is to show that incorporation and trapping is not the only route to formation of bipartite stems: in the Nakh-Dagestanian language family of the eastern Caucasus, interposition appears to be not the accidental byproduct of incorporation but the driving force behind formation of bipartites. Nakh-Daghestanian, as will be shown here, is both a geographical and a typological outlier from the North American languages with bipartite verb stems.

A tree for the Nakh-Daghestanian (or East Caucasian) family, which is about 8000 years old, is in (1). The Nakh branch, though phylogenetically half of the tree, is internally shallow and contains only three languages (Chechen, Ingush, and Batsbi); the Daghestanian branch is internally old and diverse, with about thirty daughter languages.

2. Gender agreement in Nakh-Daghestanian languages

(2)-(3) below show Ingush (Nakh branch), in which gender agreement is maximally transparent: initially or following the deictic prefix, the verb root has a gender prefix which has no allomorphy. (Gender markers are underlined.)[1]

(1) Nakh-Daghestanian (East Caucasian) family tree with schematic coordinates. (Down to the early Middle Ages, branches only; dashed line: alternate placement of Xinalug; triangle: considerable internal structure for the inner core of the Lezgian branch); D = Daghestanian branch, L = Lezgian branch)



(2) Ingush	B	jett	wa-bizhar	'the cow lay down'
	D	zhwali	wa-dizhar	'the dog lay down'
	V	Ahwmad	wa-vizhar	'Ahmed lay down'
	J	Peat'amat	wa-jizhar	'Fatima lay down'
			down-GENDER.lay	

(3)	B	suona	jett	duqa	bieza	'I like (the) cow'
	D	suona	zizaazh	duqa	dieza	'I like (the) flowers'
	V	suona	Ahwmad	duqa	vieza	'I like/love Ahmed'
	J	suona	Peat'amat	duqa	jieza	'I like/love Fatima'
		I.DAT		a lot	GENDER.like	

Hunzib (Tsezic; van den Berg 1995:79-80), shown in (4)-(6), is somewhat less transparent: one of the gender categories has zero marking; the prefixes have phonologically conditioned allomorphy as shown in (4)-(5), where a stem with a nasal vowel nasalizes the *b-* and *r-* prefixes; some verbs, as shown in (6), have vowel alternation rather than prefixation as their agreement marking.

In both of these languages, as in most of the family, gender is a partial category: some verbs take gender agreement and others do not.

(4) Hunzib	1	ozhe	yq'.lə-r	'the boy grew up'
	2	kid	j-yq'.lə-r	'the girl grew up'
	3	kãy	j-yq'.lə-r	'the berries grew'
	4	gudo	b-yq'.lə-r	'the hen grew up'

[1] Here and below, Chechen, Ingush, and Batsbi examples are written in the all-Latin no-diacritics transcription used in the UC Berkeley Chechen and Ingush projects. *ch, sh* as in English; *ss, tt,* etc. geminate/ fortis; *aa*, etc. long; *w* = pharyngealization or pharyngeal segment.

	5	q'əra	r-yq'.lə-r	'the child grew up'
(5)	1	ozhe	ũcu-r	'X stole a boy'
	2	kid	j-ũcu-r	'X stole a girl'
	3	kãy	j-ũcu-r	'X stole berries'
	4	gudo	m-ucu-r	'X stole a hen'
	5	q'əra	n-ucu-r	'X stole a child'

(6)	1	iju-l	ozhe	gəl-ər	'mother put the boy down'
	2	iju-l	kid	gil-er	'mother put the girl down'
	3	iju-l	celu	gil-er	'mother put the drum down'
	4	iju-l	baba	gul-ur	'mother put the bread down'
	5	iju-l	q'əra	gul-ur	'mother put the child down'
		mother-ERG		put down	

3. Nakh-Daghestanian bipartite verb stems

In the modern Nakh-Daghestanian languages, the verb root shape is mostly CV(R)C. For the protolanguage, however, the situation seems to have been different: for no cognate set in my database can the first consonant be reconstructed, while the second one shows regular or near-regular correspondences (Nichols 2003). Instead of a reconstructible first consonant there is usually gender prefixation and/or other, usually non-cognate, elements in the daughter languages. (7) is a representative cognate set showing gender prefixation in most daughter languages but infixation in Archi. ("=" indicates gender marker boundary, following the convention of Kibrik and Kodzasov 1988, 1990.)

(7) *=Vt 'leave/let'

Chechen		=it
Avar (Chadakolob)		=Vta
Andi		=eta
Lak	(q'a) +	=it
Dargi		=a(r)t
Archi		a=ti

3.1. Lezgian.

In most Lezgian languages, many or most verbs are bipartite. Gender and imperfective aspect are infixed to bipartite stems (immediately before the final element) but prefixed to others. Tsakhur, shown in (8), is typical.

(8) Tsakhur bipartite verbs in aorist tense, showing infixation of gender (Dobrushina 1999:85; qq = geminate, y = high back unrounded vowel, X = uvular)

Gender	'hold' {a=q}	'hang' {giwa=X}
1	aqqy	giwarXyn
2	ajqy	giwarXyn
3	awqu	giwapXyn
4	aqqy	giwatXyn

Three Lezgian languages have lost gender: Lezgi, Agul, and Udi. Without gender infixation, there is no way of distinguishing bipartite verbs from ordinary compounds like those in (9).

(9) Lezgi compound verbs (Haspelmath 1993:171)

aw-atun	'fall out'	aw-udun	'take out'
al-atun	'fall off'	al-udun	'take off'
ak-atun	'fall under'	ak-udun	'take away'
hal-tun	'meet'	hal-dun	'cover, put on'

3.2. Dargi. Kubachi Dargi has an large set of preverbs in two position slots, but these form compound rather than bipartite verbs. Gender agreement is initial on the verb root; some preverbs also have gender agreement, prefixal for some and suffixal for others. Most preverbs are transparently derived from postpositions. In other Dargi varieties, preverb-root combinations are lexicalized to the point that theycan be considered bipartite stems: van den Berg, in press.

(10) Kubachi Dargi (Magometov 1963:76)

b=e:n - ka - b=iššši-j 'insert, put in' (gender class III)
w=e:n - ka - w=iššši-j 'go in' (gender class I)
GENDER=in-down-GENDER=go-INF

3.3. Lak. Gender is prefixal except that a few verbs also infix gender and are thus bipartite.

(11) Lak gender agreement (Zhirkov 1955:93, [ ] = constructed; Xaidakov 1962:418)

Simple stem:	bucara	[ducara]	ucara	'brings'
Bipartite stem:	buvna	durna	uvna	'went'

3.4. Avar-Andic and Tsezic. Gender is mostly prefixal except that a very few verbs infix gender and are therefore bipartite.

(12) Avar-Andic 'wash, launder' (Gudava 1959:197, his set no. 218). Boundary "=" added. *b-* is traditional citation form for gender prefix.

	Stem	Past tense
Avar	čur-	čurana
Andi	a=b=č-	abčo
Botlikh	m=uč-	muča
Chamali	b=uš-	bušã
Tindi	b=uč-	bučõ
Bagvali	b=učw-	bučã
Karati	b=ičw-	bičwã (Anchikh dial.: m=učw-, muča)
Akhvakh	č-	čari

(13) Avar-Andic 'weave' (Gudava 1959:142, set no. 132.)

	Stem	Past Tense	
Avar	b=ess-	b=essana	
Andi	b=ešš-	b=ešši	
Botlikh	šš-	arxaššu	
Godoberi	hišš-	hišši	
Chamali	iss-	issi	
Tinti	išš-	iššo	
Bagvali	eš-	ešo	
Karati	kešš-	keššẽ	
"	keršš-	kerššĩ	(Tokita dial.)
Akhvakh	goss-	gosseri	
"	k'anss-	k'anssira	(Tljabon dial.)

3.5. Nakh.

In the Nakh languages, gender is exclusively prefixal in roots and never infixed; see again (2)-(3). Uniquely in the family, the Nakh branch has variable initials in a number of verbs, where gender prefixation is just one of the different initials and the others are not segmentable morphemes in Nakh or Nakh-Daghestanian (though some of the same consonants recur as initials in different sets). The vowels and ablaut classes also vary, but this is common in the Nakh branch. Examples are in (14)-(17). An abstract meaning for the verb root is suggested for each set. ? = connection uncertain.

	Ingush	Chechen	Batsbi	gloss
(14)	tassar	tasar	tasar	strew, scatter, sprinkle
	qossar	qossar	qosar	throw, cast
?	moassa+d=ar	--		splash; pour water out

Semantics: *-ss / -s* motion of thrown, tossed, or poured mass or particles

(15)	ottar	ottar		stand up
	--	hottar	--	stand up
	laattar	laattar	lattar	be standing
	d=ottar	d=ottar	d=ottar	stack; pour
	ghattar	ghaattar	--	fly away, take off
	--	--	qettar	stand up.ITER
		cf. litta 'haystack'		

Semantic analysis: *-tt* verticality of motion or stance

(16)	allar	willar		lie, be lying
	d=ollar	d=ollar	d=ollar	insert, enclose, bury
	d=ylla	d=illar	d=illar	lay foundation; plate
	qollar	qollar	qollar	cast, cover
	tyllar	tillar	tillar	put on top; name
?	d=aallar	d=aallar		be tightly contained, inserted

?	ollar	ollar	ollar	hang; string up, string together

Semantics: *-ll* contact over side, plane, surface

(17)	d=aagar	d=aagar	d=ak'	burn
	sagar, sog	--		set afire
		sagar, sega		flash, shine
			ak'+d=ar	burn
		qaagar+d=uolar		scorch, burn and stick to pan

Semantics: **-k'* burn

These are cognate sets in which, within Nakh, the second consonant reconstructs but the first does not. They must be the partly or mostly intact remains of Proto-Nakh sets that may once have been larger. At least back to Pre-Proto-Nakh, there is no reconstructible evidence that they were ever semantically more transparent.

In addition, the Nakh languages have inflectional and derivational prefixation of deictic, directional, and locative elements. When these are prefixed to a gender-agreeing verb root, the result is a trapped gender prefix (much as in (10)). Ingush examples:

(18)	hwa-aara-v=ealar	dwa-chy-d=ellar
	here-out-GENDER=went	there-in-GENDER=inserted
	'he came out (toward speaker)'	'inserted (e.g. key in lock)'

4. Comparative grammar of Nakh-Daghestanian bipartites

4.1. Verb root structure.

Why is only the second consonant reconstructible for Proto-Nakh-Daghestanian (PND) verb roots? There are five synchronic patterns that are responsible for this:

(a) The first consonant slot is occupied by a gender prefix in most branches. An example is (7) above. In (20) below Agul *f-* is a variant of a frozen gender marker (Agul has lost gender) and not the regular reflex of any PND consonant.

(b) Occasionally, the first consonant is a reduplicate of the second. There are sporadic examples in Lezgian and Lak; see (19) and (21) below.

(c) Most of the languages have a few monoconsonantal verbs. The verb 'see' is monoconsonantal in a number of daughter languages; see (19) below.[2]

(d) There are occasional other initials. *l seems to recur in more than one branch; it may or may not be the same element as the *R of the canon form *(C)V(R)C. In (20) below, Chechen and Ingush have *l-* in one form of the root, no initial in the other; Chamali and Agul have no sonorant imediately before the reflex of **-c*; Dargi has a sonorant not connected to gender; and Tabassaran and Archi have gender markers immediately before the reflex of **-c*. In (23), Avar and Dargi have C=VRC roots where the sonorant is distinct from the gender agreement; Agul

[2] Two of the monoconsonantal Nakh verbs consist of nothing but a gender prefix and a tense suffix: Ingush pres. *d=u*, past *d=yr* (Chechen *d=o, d=ira*) 'do' and pres. *d=y*, past *d=ar* (Chechen *d=u, d=ara*) 'be'. They belong to different conjugation classes but have no segmental lexical identity.

has a sonorant and no gender; Archi has a second gender marker in this slot; and the other languages have no sonorant. Apart from these possible connections of *l-* to *(R) and gender, there are no consistencies in initial elements.

(e A bipartite initial occupies the first consonant slot. This is especially common in the Lezgian branch:

> The *preverb* [=my *initial* – JN] is the initial element of simple roots. It can be segmented off as a separate element because the gender/number slot comes between it and the root consonant. Though preverbs in origin, synchronically these elements cannot be considered productive derivational morphemes because no regular meaning can be ascribed to them. (Dobrushina 1999:58, on T͡sakhur (Lezgian).)

(19)-(23) are cognate sets showing various discrepant initials in Nakh-Daghestanian. (‡ = source other than Kibrik & Kodzasov 1988; ? = connection to this cognate set uncertain. Glosses are identical to the first one unless otherwise indicated.)

(19)	Chechen, Ingush	gar	'see'	<u>initial</u>: none
	Andi	haGwo		h-
	Bezhta	=egaa		gender
	Lak	kkwakkw		reduplicate
	Dargi	ʕul + =irg (ʕul 'eye')		gender
	Agul	agw		none
	Archi	=akkwa		gender

(20)	Chechen, Ingush	laacar	'catch, capture'	<u>initial</u>: *l-
		iecar	'take, buy'	none
	? Lak	las	'carry, bring'	*l-
	? Lak	=u=c	id.	gender
	Chamali	=Vci	'take, hold'	gender
	Dargi (Chirag)	=u(r)c	id.	gender
	Tabassaran	=i=s	id.	gender
	Agul (Richa)	fac	id.	f- < *gender
	Archi	=sa	id.	gender

(21)	Chechen, Ingush	laq'	'go dry, dry up'	<u>initial</u>: *l-
	Avar	b=aqq'waze	‡	gender
	Andi	b=eqq'uda	‡	gender
	Lak	q'aq'an	‡	reduplicate
	Lezghi	q'urun	‡	none
	Archi	q'uras	‡	none

(22)	Chechen, Ingush	d=ouzar	'know, recognize'	<u>initial</u>: gender
	Lak	=uvč'	'understand'	gender
	Avar (Chadakolob)	=ic'	'understand'	none
	Andi	c'indi	'know'	none
	Rutul	ac'	'know'	none
	Tsakhur	=c'a=xes	'know' [my segmentation -- JN]	none

(23)	Chechen, Ingush	d=ustar 'measure'	<u>initial</u>:	gender
	Avar (Chadakolob)	=orš(n)		gender
	Andi	=ason		gender
	Bezhta	=ašah		gender
	Lak	=uc		gender
	Dargi	=unc		gender
	Agul	alc		none
	Archi	=a=sin		gender (2x)

Since monoconsonantal verbs are very few (synchronically and for the protolanguage), initials are usually present but discrepant, and no set of prefixes can be reconstructed, the most conservative and least committal reconstruction is to posit that PND had many bipartite verb stems whose internal segmentation may already have been less than transparent.

4.2. Inflection.

It should be noted that the gender markers, as a set, seem to reconstruct for PND, but they seem to have been originally head class markers (markers on the nouns themselves rather than agreement markers on the verbs: Nichols in press, 1990, work in progress; for head class see Evans 1997, Evans et al. 2002, the latter using the term *morphological class*). Their use in agreement seems to be a secondary, though still very early, post-Nakh-Daghestanian innovation that arose in the southeast of the family's range, in the Lezgian branch, and affected the northwestern branches, Nakh and Avar-Andic (±Tsezic) later. Where it arose earlier it is least transparent, most often infixal, and found in most or all verbs; where it spread later it is more transparent and affects fewer verbs. Head class marking is evident now in Nakh and Xinalug (Nichols in press, 2003) and probably in no other branch.

Interposing or infixation is widespread, though usually not productive, in the modern Nakh-Daghestanian languages. Gender prefixation is found in all branches, though with very different frequencies. Examples are in (24)-(25).

(24) Interposed and/or infixed elements in modern Nakh-Daghestanian branches

Nakh	Pluractional categories: multiple argument labial infix
	Iterative **e/*i* ablaut
Avar-Andic	Gender (a very few verbs)
Tsezic	Pluractional (infix);
	Gender (ablaut) (21 verbs: van den Berg 1995:80)
Lak	Gender (a few verbs)
Dargi	Pluractional (*r/l* infix; ablaut)
Lezgian	**i* ablaut and **l* (Schulze) or /r/ (Alekseev) infix (durative)
	Gender (ablaut or infix) (most verbs)

(25) Inflectional prefixation of verbs in Nakh-Daghestanian branches.

Nakh	Gender (about 30% of verb roots)
Avar-Andic	Gender (about half of verb roots)
Tsezic	Gender (about half of verb roots)
Lak	Gender (most verb roots)
Dargi	Gender (many verb roots)
Lezgian	Gender (non-infixing verbs; most verbs are infixing)

Variable initial consonants as in (14)-(17) are found only in the Nakh branch, unless Dargi verb pairs in which the perfective member has gender prefixation and the imperfective lacks it (Magometov 1963:165ff. for Kubachi) also belong here.

To summarize: Infixation of pluractional categories of some kind, including at least an *i infix, appears to be reconstructible for PND. Gender prefixation and infixation may be secondary (though still very old). Originally root-internal sonants of the *CV(R)C pattern are likely to have been recruited to gender agreement, and this may also have been the source of the sonorants involved in pluractional infixation in some branches. The variable initial consonants now found only in Nakh are reconstructible, at least as a pattern: if the variation *within* languages is found only in Nakh, there is evidence for it *between* the other branches (see (19)-(23)).

4.3. The set of bipartites. A likely PND bipartite stem is shown in (26).

(26) A reconstructed bipartite stem: **(C)=VRdz* 'stand, stand up'. (Some of the Daghestanian initials may be recently added preverbs, but probably not all.)

PND:	* Ø- / h-	* uvular	* l-	other initials
Nakh:				
Chechen	hott- 'stand up'	ghaatt- 'fly off'	laatt- 'stand'	d.ott- 'stack'
Ingush	ott-	ghatt-	laatt	d.ott-
Batsbi	ott-	qett- id.ITER	latt-	d.ott-
Daghestanian:				
Avar *	c'(Vl) 'stand'			
Andi	hirc'i 'stand up'			
Chamali *	hinc'i 'stand up'			
Bezhta *		=aghic' 'stand up'		
Lak	=iz 'stand up'			
Dargi	he-=i(r)cc 'stand up'			ka-=Vrcc 'stand'
Lezgi		aqqw[-]azun 'stand'		
Tabassaran *		GV-udzw 'stand up'		
Agul (Burschag)	gha-zw-a 'stand up'			
Rutul			lu=zw 'stand up'	
Tsakhur			ulyo=z-r	
Budukh		Gu=zVr		
Archi				=o=ci 'stand'

* Dialects cited: Chadakolob Avar, Gigatl Chamali, Tljadal Bezhta, northern Tabassaran.

PND bipartite verbs may have been a closed class, as indicated by the fact that pluractional ablaut or infixation generally affects a smallish closed class of verbs in the daughter languages, and a minority of the Nakh verb roots show variable initials. (In most of the daughter languages, verb roots themselves are a closed class; new verbs are formed by directional/locative prefixation in the minority of languages that have it, and in all languages by compounding with auxiliaries to form phrasal verbs.) No non-bipartite *CVC or *CVRC stems appear to be reconstructible, so apparently the alternative to bipartite structure was not unipartite *CV(R)C structure (as it is in several daughter branches) but monoconsonantal *(V)C structure.

While bipartite structure in the abstract can be reconstructed for PND, specific bipartite verbs are harder to reconstruct. (20)-(23) above are likely to have been bipartite. (26) above is the clearest case of a bipartite PND verb. Otherwise, the members of the closed bipartite sets of Nakh, Avar-Andic-Tsezic, and Lak do not coincide. This suggests that each such set is a shrunken residue of the PND bipartites, which must have been a larger (though still probably closed) set.

Thus, the only aspect of bipartite structure that can be firmly reconstructed is the variable initial consonants, though they are clearly in evidence in only one branch.

5. More recent kinds of bipartitivity

There are more recent ways in which verb roots consist, or come to consist, of two pieces in various daughter languages.

5.1. Reduplication.

For ancient reduplication see (19), (21). In Rutul (Lezgian), gender is infixal for CV(R)C stems but prefixal for reduplicated stems (Alekseev 1985:87), so this reduplication must be secondary.

(27)	Reduplicated:	ra=q'alq'as, va=q'alq'as	'tremble'
	CVC:	ki=r=kas, ki=v=kas	'resemble'

Rutul also has various kinds of inflectional reduplication not illustrated here.

5.2. Chechen-Ingush whole-root reduplication.

The chaining enclitic =*'a* must be interposed between the verb root and the preverbal element (directional prefix, first element of compound verb, direct object), and if there is no such host the verb root is reduplicated to provide one (Peterson 2001, Conathan & Good 2000). Ingush examples showing various hosts are in (28)-(31); the clitic is in boldface.

(28) First element of verb:
yz bwarjg=**'a** veina
3s eye **&** V.see.CV
'saw him and ...', 'having seen him, ...'

(29) Direct object:
mashen=**'a** iicaa
car **&** buy.CV
'bought a car and ...'

(30) Preverb:
chy=**'a** veanna
in **&** V.go.CV
'went in and ...'

(31) Reduplicate :
Yz kinashjka hwa='a iicaa, diisha=**'a** diishaa, waqessar cuo
this book DX & buy D.read **&** D.read.CV down-throw.WP 3s.ERG
'He bought, read, and threw out the book'

5.3. Udi endoclisis.

In Udi (Harris 2000, 2002), person markers are clitics placed at the right edge of the focused element. (The person-number clitic is an innovation in Udi.) The person marker appears between the parts of a complex stem (28), between the two elements of a bipartite stem ((29)), or – if the stem is simplex – before its last consonant (30)-(31). (Σ = first part of bipartite stem.)

(28) me pasčagh-en eč-es-**ne**-st'a ... kul
this king-ERG bring-INF-**3s**-CAUS earth
'This king has earth brought ...' (Harris 2002:122)

(29) nana-n bạgha-**ne**-b-e p'ạ ačik'alshey
mother-ERG find-**3s**-do-AOR two toy
'Mother found two toys' (p. 122)

(30) kaghuz-ax a-z-q'-e
letter-DAT Σ-1s-receive-AOR
'I received the letter' (p. 125)

(31) similarly: a-ne-q'-sa u-ne-k-sa
Σ-1s-receive-PRES Σ-1s-eat-PRES
'receives' 'eats' (p. 127)

5.4. Phrasal and compound verbs.

In most of the daughter branches simple verb roots are a closed class, so phrasal predicates and compound verbs of various kinds are numerous.

6. Structural explanation

There are several unusual, even mysterious, factors in the structure and history of Nakh-Daghestanian bipartites. Gender infixation is their most frequent and probably best-attested identifying mark, but it is likely to be secondary; discrepant initials probably reflect the inherited state, but they are found only in a handful of verbs in one branch. That is, gender infixation is found only in the Daghestanian half of the family and variable initials only in the Nakh half; the daughter languages share the commitment to bipartite structure, but realize it in different ways. The origin of gender agreement seems to have involved head class markers being moved into or copied inside of verb roots. Pluractional infixation is found in both halves of the family, but is likely to also be secondary, independently formed by recruitment of various root-internal elements to number inflection. Secondary, sometimes very recent, bipartite properties are found in various branches of the family. The two most recent ones are quite unusual: the Chechen-Ingush coordinating/chaining enclitic is so far the world's only firm example of a clitic which is positioned relative to the final element in its domain, precedes that, and is proclitic to the word before it (Peterson 2001; a Type 5 clitic in the typology of Klavans 1985), and the no less curious Udi endoclisis inserts a clitic into an indivisible word (Harris 2000, 2002).

Strange though this set of developments is, they can be subsumed under a panchronic generalization that covers both ancient and secondary bipartite phenomena in Nakh-Daghestanian languages: Elements such as clitics, and others that are positioned relative to domain edges, are always placed before the final element of the domain, in a reverse Wackernagel position. Furthermore, what constitutes an "element" of the domain for purposes of this positioning must be of the same kind as the element to be positioned. Thus, in most of the daughter branches, focus position in the verb phrase or clause is immediate preverbal position; since the verb is VP-final and in nearly all of the languages clause-final, this amounts to placing the focused word immediately before the domain-final word. A clitic in this position is proclitic to the last word, e.g. the verb in the verb phrase as in (32), or enclitic to the word before that (the position described by Peterson 2001), as in (33)-(34).

(32) Ingush dika cy= xouzh
well NEG know.CONVERB
'not knowing well', 'since (s/he) doesn't know well...'

(33) Ingush cynga axcha ='a d=anna
him.DAT money & GENDER=give.CONVERB
'gave him money and ...', 'having given him money, ...'

(34) Ingush bwarjg ='a v=eina
eye & GENDER=see.CONVERB
'saw him and...', 'having seen him, ...'

An affix is placed before the last suffix of its word. Thus, when the negative proclitic became an affix in Ingush, it moved to before the tense ending (PST = past stem formative, PAST = witnessed past tense ending):

(35) Proto-Nakh *co v=aal-in-ra
NEG V=go-PST-PAST

Pre-Ingush *vaal-in-**c**-ira

Ingush veal-an-**dz**-ar
go-PST-**NEG**-PAST
'didn't depart'

Neither Chechen nor Batsbi suffixes the negative: cognate to the Ingush form above is Chechen *ca veelira.* Batsbi does, however, suffix the reported marker in the same position: {tet'-i-no-ra} /tit'nor/ [cut-PST-REPORTED-PAST] '(he/she) cut': Holisky 1994:180.

Finally, a segment is positioned before the last segment of the stem. Examples are the pluractional infixes as in (36), and gender infixation in all languages that have it.

(36) Labial pluractional in Nakh:

Batsbi	tasar	pl.	tepsar	'throw'
	qollar		qeblar	'cast'

This set of positioning tendencies can be stated as in (37).

(37) **Nakh-Daghestanian strong interposition rule:** Any floating or clitic item preferably goes before the domain-final element at its own level: word before word, affix before affix, segment before segment. Gender agreement markers and other monosegmental formatives can be treated as segments.

This principle explains the positioning of everything except the Udi endoclitics, for which the only oddity is that they are treated as though they were monosegmental (when most are bisegmental) and positioned before a final segment rather than before a word, as clitics should be.[3]

References

Alekseev, M. E. 1988. *Sravnitel'no-istoricheskaja morfologija avaro-andijskix jazykov.* Moscow: Nauka.

----. 1985. *Voprosy sravnitel'no-istoricheskoj grammatiki lezginskix jazykov.* Moscow: Nauka.

Conathan, Lisa, and Jeff Good. 2000. Morphosyntactic reduplication in Chechen and Ingush. *CLS.*

DeLancey, Scott. 1996. The bipartite stem belt: Disentangling areal and genetic correspondences. David Librik and Roxane Beeler, eds., *BLS 22 Special Session on Historical Issues in Native American Languages,* 37-54. Berkeley: BLS.

Dobrushina, Nina. 1999. Glagol. 58-85 in A. E. Kibrik, ed., *Èlementy caxurskogo jazyka v tipologicheskom osveschenii.* Moscow: MGU.

Evans, Nicholas. 1997. Head class and agreement class in the Mayali dialect chain. In *Nominal classification in Aboriginal Australia,* eds. Mark Harvey and Nicholas Reid, 105-146. Amsterdam: Benjamins.

Evans, Nicholas, Brown, Dunstan, and Corbett, Greville. 2002. The semantics of gender in Mayali: Partially parallel systems and formal implementation. *Language* 78:111-155.

Gudava, T. E. 1959. *Sravnitel'nyj analiz glagol'nyx osnov v avarskom i andijskix jazykax.* Makhachkala: IIJaL Dag. Fil. AN SSSR.

Harris, Alice C. 2000. Where in the word is the Udi clitic? *Language* 76:3. 593-616.

----. 2002. *Endoclisis and the Origins of Udi Syntax.* Oxford University Press.

Haspelmath, Martin. 1993. *A Grammar of Lezgian.* Berlin: Mouton de Gruyter.

Holisky, Dee Ann, and Rusudan Gagua. 1994. Tsova-Tush (Batsbi). Rieks Smeets, ed., *The Indigenous Languages of the Caucasus, vol. 4: North East Caucasian Languages, part 2,* 147-212. Delmar, NY: Caravan Books.

Jacobsen, William H., Jr. 1980a. Washo bipartite verb stems. Kathryn Klar, Margaret Langdon, and Shirley Silver, eds., *American Indian and Indoeuropean Studies: Papers in Honor of Madison S. Beeler,* 85-99. Berlin: Mouton de Gruyter.

Kibrik, A. E., and S. V. Kodzasov. 1988. *Sopostavitel'noe izuchenie dagestanskix jazykov: Glagol.* Moscow: Moscow University.

----. 1990. *Sopostavitel'noe izuchenie dagestanskix jazykov: Imja. Fonetika.* Moscow: Moscow University.

Klavans, Judith. 1985. The independence of syntax and phonology in cliticization. *Language* 61.95-120.

Magometov, A. A. 1963. *Kubachinskij jazyk.* Tbilisi: Georgian Academy of Sciences.

Nichols, Johanna. 2003. The Nakh-Daghestanian consonant correspondences. Dee Ann Holisky and Kevin Tuite, eds., *Current trends in Caucasian, East*

[3] This paper is based in part on work supported by the National Science Foundation under Grant 9616448.

European, and Inner Asian linguistics: Papers in honor of Howard I. Aronson. Philadelphia: Benjamins.
----. In press. Chechen (with notes on Ingush). Alan S. Kaye, ed., *Morphologies of Asia and Africa (including the Caucasus).* Winona Lake, IN: Eisenbrauns.
----. 1990. The Nakh evidence for the history of gender in Nakh-Daghestanian. H. I. Aronson, ed., *The Non-Slavic Languages of the USSR: Linguistic Studies,* 158-75. Chicago: Chicago Linguistic Society (University of Chicago).
Peterson, David A. 2000. Ingush *ʔa*: the elusive Type 5 clitic? *Language* 77.144-55.
Schulze, Wolfgang. 1998. *Person, Klasse, Kongruenz. Fragmente einer Kategorial-typologie des einfachen Satzes in den ostkaukasischen Sprachen. Vol. 1: Die Grundlagen, part 1.* Munich: Lincom Europa.
van den Berg, Helma. 1995. *A Grammar of Hunzib (with texts and lexicon).* Munich: Lincom Europa.
----. In press. Spatial prefixes in Dargi (East Caucasian). F. Kiefer, ed., Proceedings of Tenth Morphology Meeting. *Folia Linguistica Hungaric* special issue.

Establishing a Community of Practice on the Internet: Linguistic Behavior in Online Japanese Communication*

YUKIKO NISHIMURA
Toyo Gakuen University, Japan

1. Introduction

In this paper I will analyze linguistic behavior in online communication in Japanese. Specifically, I will describe certain unique uses of language on the Bulletin Board Systems at a particular Japanese website called "*ni channeru*", or "Channel Two". This website was created in May 1999 by an individual named Hiroyuki Nishimura, age 26, for users to share funny stories, jokes, opinions and so on. This website contrasts with other BBS websites where more serious exchanges of useful information take place.

I will apply to this virtual community the theoretical framework of the "community of practice", proposed by Wenger (1998), who says, "People congregate in virtual spaces and develop shared ways of pursuing their common interests" (7). In this framework, there are three requirements for a community of practice: (1) a joint enterprise, (2) mutual engagement, and (3) a shared repertoire. First, the users' joint enterprise is to construct this website; they do so by sending and reading messages. Second, they are mutually engaged in discussing topics of interest. And third, concerning a shared repertoire, users have developed a fairly large number of unique language forms and playful discourse practices, which will be the focus of this paper.

In the Internet community of Channel 2 users, there are clearly members who are more or less experienced in website activities. Thus Wenger's notion of core versus peripheral membership works well to describe the participants in this community of practice. The oldest member of the community is considered to be the creator, Hiroyuki, the only publicly known figure. Next come the voluntary helpers who assist Hiroyuki in maintaining the computer servers, who have the authority to delete messages; because they know the website rules and guidelines, they sometimes educate other users. Next are certain users who consistently use the same handle names, lead discussions, post messages

*This research was conducted while I was a visiting researcher in the Department of Linguistics, University of California, Santa Barbara. I am grateful to Patricia M. Clancy and Mary Bucholtz for their comments on earlier versions of this paper. Any errors that may remain are my responsibility.

frequently, and are considered to be active users. These groups of users form the core membership of the community. Then come the vast majority of anonymous senders, and surrounding them are even larger numbers of users who only view messages; these users can be regarded as peripheral members. About 15 to 20 percent of the users who visit the site and view messages actually post messages. Since a sender on one occasion can also be a viewer on others, this is an approximate categorization of the membership.

The topics discussed on the website cover almost every conceivable subject, categorized into over three hundred boards and further subcategorized into two to three hundred threads per board. The estimated users' age range is from preteens to fifty- or sixty-year-olds. There is no way of obtaining accurate user profiles, such as age, gender, occupation, geographical residence, and so on, because of the strictly guaranteed anonymity on the website, though such information may sometimes appear in their posted messages.

Before turning to the data, I will explain how computer-mediated communication takes place in Japanese. There are four kinds of scripts in standard Japanese orthography, and this has an important effect on the users' behavior in Internet communication.

2. Computer-Mediated Communication in Japanese

Interpersonal and informal CMC differs from face-to-face communication in that auditory and paralinguistic information is lacking; furthermore, entering a message on the keyboard takes more time. Users therefore employ emoticons or face marks to show facial expressions, such as ☺, and acronyms (*e.g. LOL* for *laugh out loud*) to speed up typing. Like CMC in other languages, Japanese also uses these means (See Nishimura (to appear) for a more detailed description). However, because the scripts used in Japanese differ from those used in English and other languages, there are major differences in the kinds of realization, interpretation, and interactional effects that can be achieved.

2.1. Four kinds of scripts in Japanese orthography

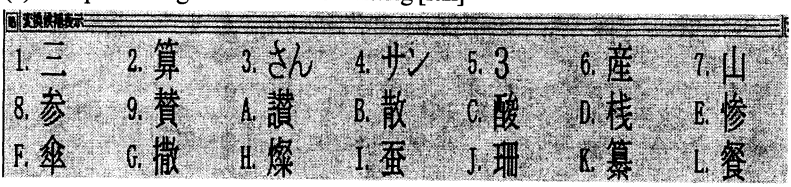
In standard orthography, the Japanese language employs four kinds of scripts, each with its own functions and uses in the Japanese writing system: (1) *Hiragana,* a syllabary, used for certain native nouns, and for grammatical markers, such as postpositions and verb inflections; (2) *Katakana,* a syllabary, typically used for writing foreign (especially Western) names, loan words, and so on; (3) *Romaji*, Roman letters, used to transliterate personal and place names; (4) *Kanji,* characters originally borrowed from Chinese. Each character represents a particular meaning, but can be pronounced in two completely different ways. For example, for the character that means *mountain*, one pronunciation is [jama], and the other one is [san] or [zan], depending on the phonetic context.

Because of this variety in scripts, it is possible in Japanese to write the same word in four different ways: e.g. to express "*nihon*" 'Japan', the *hiragana* representation is "にほん"; *katakana*, "ニホン"; *Romaji*, "*nihon*"; and *kanji*, "日本". One of the reasons for using *kanji* instead of *hiragana* and *katakana* is to avoid ambiguity. Since there are so many homophones in Japanese in general, *nihon* 'Japan' in the representations other than *kanji* could also mean "two long cylinder-like objects, but the *kanji* representation, 日本, which uses ideographs, is unambiguous.

2.2. Word processing in Japanese

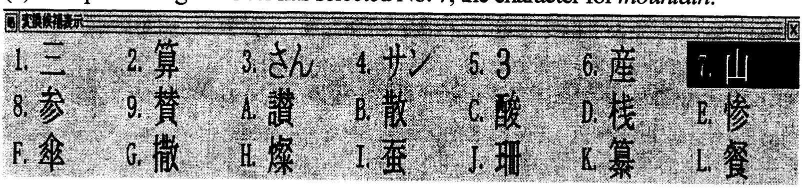
To write in Japanese on the computer, users must enter words in *romaji* in the Japanese input mode, and then the conversion software changes the *romaji* into the *hiragana* syllabary. Next, the system automatically shows the users' most recently used *kanji* corresponding to the *hiragana*. In (1) below, we see the list of *kanji* for [san], which was typed in *Romaji* as *san*: In this list, No. 1 is the Chinese numeral character for *three*; No. 2 is the *kanji* meaning 'computation'; No. 3, *hiragana*; No. 4, *katakana*; No. 5, the Arabic numeral; No. 6, the *kanji* meaning 'production'; No. 7, the *kanji* meaning 'mountain', and so on, all of which are pronounced [san].

(1) Computer Image 1: Choices for writing [san]



When the first *kanji* shown by the software is not the intended character, users must press the space key to make the system show a list of all the other *kanji* characters of other representations that have the same pronunciation. To choose one of these options the user presses "enter" to select the desired character. (2) below shows the user's choice.

(2) Computer image 2: User has selected No. 7, the character for *mountain.*



Users can also choose *hiragana* (No.3) or *katakana* (No. 4) at this stage after the list is shown. Obviously, users might choose the wrong *kanji*, either by mistake or deliberately, as we shall see.

Because each *kanji* character has a meaning, Japanese users can use *kanji* characters for the same purpose as acronyms and emoticons. Observe the sentences in (3) through (5), which are constructed for the purpose of illustrating the functions of *kanji* in online messages:

(3) そんな こと って、 ある の！ (笑)
sonna koto tte, aru no wara-
such thing QUOTE exist/happen SFP laugh STEM
'Could such a thing happen?!' '*kanji* for LAUGHTER'

(4) そんなことって、ある の！ (涙)
namida '*kanji* for TEARS'

(5) そんなことって、ある の！ (喜)
yorokobi '*kanji* for JOY '

In these sentences, the verbal message is identical; the differences are in the *kanji* in parentheses at the end of each message. This practice of adding a *kanji* character in parentheses at the end of a message, especially the character for *laughter*, is very common in Japanese CMC.[1] In addition, however, Channel 2 users treat *kanji* in very innovative and unique ways that ignore the characters' meaning. I turn to these uses now.

3. Unique linguistic practices in Channel 2

The data I analyze here come from three sources: first, the "Glossary of Channel 2 Terms" available online; second, its printed version, *Channel 2 Dictionary;* and third, actual messages posted on Channel 2. Specifically, this study identifies uses that deviate from the standard, and analyzes them regarding their purposes, origins, and process of formation.

I have classified the unique linguistic practices in Channel 2 into three major groups: first, non-standard uses of *kanji* as phonetic representations; second, play with script shape similarity; and third, errors that have become conventionalized on this website.

3.1. *Kanji* as phonetic representation: *kanji* punning

In the first major group, *kanji* as phonetic representation, the meaning of the *kanji* is ignored and it is used as a phonetic representation for words that have the same pronunciation but different meanings. I call this phenomenon "*kanji* punning", and it is fairly common on this website. There are a huge number of homophones in Japanese, and in Channel 2, the users apparently enjoy exploring the tremendous gap between the intended and the literal meanings of their puns. The major types of *kanji* punning are described in Sections 3.1.1-3.1.3 below.

3.1.1. *Kanji* for abbreviated loan words

Three examples of the first type of phonetic use, *kanji* for abbreviated loan words, are given in Table 1 below. What is unconventional is the use of *kanji* to represent loan words.

[1] Other *kanji* characters that can function this way include: (謎) *nazo* 'puzzlement', (照) *tere* 'shyness', (恥) *haji* 'shame', (爆) *baku* 'explosion, burst', (汗) *ase* 'perspiration', (死) *shi* 'death', (苦笑) *kusyou* 'wry smile', (鬱) *utsu* 'melancholy', and so on.

Normally, these words are represented in *katakana*, but because *katakana* representation takes more time to enter, the *kanji* shortcut is preferred. To get *katakana* representation in the case of (computer) *server*, users need to press the key 7 times until the *katakana* is finalized, but *kanji* representation requires only 5 times. Thus the primary reason for this use is to speed up typing, and at the same time, a secondary reason is to enjoy the gap between the intended and literal meaning found in the *kanji* used to express the loan word, which has already been used by other users and lexicalized on this website.

Notice, however, though representation of loan words in *kanji* takes place on this website, this phenomenon is rather limited in its application, in that *kanji* characters that have the pronunciation of the abbreviated loan words must exist. For abbreviated loan words such as *resu* formed from *resuponsu* 'response', the *kanji* character that has the reading of *resu* does not exist, and such abbreviated loan words are in most cases written in *katakana*, as レス in conventional way.

Table 1. Formation process: *kanji* for abbreviated loan words

Origin	Loan word in *Romaji*	Abbreviated *Romaji* entered	Automatic *hiragana* conversion	Automatic *kanji* conversion and literal meaning	Channel 2 meaning	Cf. Conventional orthography
a) *server*	*saabaa*	*saba*	さば *saba*	鯖 *saba* 'mackerel'	鯖 *saba* 'server'	サーバー (in *katakana*)
b) *proxy*	*purokushii*	*kushi*	くし *kushi*	串 *kushi* 'skewer'	串 *kushi* 'proxy'	プロクシー (in *katakana*)
c) *account*	*akaunto*	*aka*	あか *aka*	垢 *aka* 'dirt, filth'	垢 *aka* 'account'	アカウント (in *katakana*)

Table 1, example (a), illustrates the process by which the English word *server* comes to be written by the character for *mackerel*. *Server* becomes [sa:ba:] when it is taken into Japanese as a loan word. Then the loan word, [sa:ba:] gets shortened to [saba]. When users enter *saba* on the computer, the word is converted to *hiragana*; since the standard character corresponding to *saba* is *mackerel*, this is the character that automatically appears. In (6) below, the message writer is discussing someone's skills in setting up the computer server, and in doing so, the character that means *mackerel* is used. Table 1 also gives two additional examples of this process. Though the user could select other possibilities, the character that is shown automatically by the software is used because it is the fastest. In some cases users pick a certain character because it has become conventionalized in Channel 2 usage, even if it is inconvenient to select it over other candidates. Note also that the loan words in Table 1 relate to computer uses.

(6) そいつ 鯖 立てる 程度 の スキル も 無い …
soitsu saba tateru teido no sukiru mo nai …
that guy server/mackerel setup extent of skills even possess.NEG
Literal meaning: 'he doesn't even have enough skills to set up mackerel…
Channel 2 meaning: 'he doesn't even have enough skills to set up the server…'
(From PC News Board)

3.1.2. *Kanji* punning for speed and special connotations

The next group of phonetic uses of *kanji* is used for added speed and special connotations. In Table 2 example (a), we see the process for writing *chuubou*, a common word in Channel 2.

Table 2. Formation process for *kanji* punning : "*chuubou*" and other examples

Origin/source expression	*Romaji* after shortening/ modification	Automatic *hiragana* conversion	Automatic *kanji* conversion literal meaning	Channel 2 meaning and connotation
a) 中学生の坊主 *Chuu gaku sei no bou zu* 'middle school kid'	*chuu bou*	ちゅうぼう *chuubou*	厨房 *chuubou* 'kitchen'	厨房 *chuubou* 'immature middle school brat'
b) 小学生 *syougakusei* 'elementary school pupil'	*syougakusee*	しょうがくせえ *syougakusee*	生姜くせえ *syougakusee* 'smelling ginger'	生姜くせえ *syougakusee* 'immature elementary school brat'
c) 笑う or 笑い *warau* or *warai* 'to laugh' or 'laughter'	*wara* (stem)	わら *wara*	藁 *wara* 'rice straw'	藁 *wara* 'sneering laughter'

Chuubou comes from *chuugakusei no bouzu*, 'middle school kid', which is an ordinary expression that does not necessarily have derogatory connotations. This then gets shortened to *chuubou*. Then the *romaji* for *chuubou* is entered on the computer and converted to *hiragana*. Next, the first and only character that the system shows is the character that means *a hotel /restaurant kitchen*, as shown in Table 2. Thus on Channel 2 *kitchen* is used with derogatory connotations to mean 'immature and inconsiderate middle school brats'. Table 2 gives two additional examples of *kanji* punning for special connotations.

An example of how *chuubou* is used in an actual message is given in (7), from the *Chuubou* Board of the website. The expression *yatsura* 'guys' in (7) is also consistent with the insulting content and tone of this message.

(7)-a	ここ	に	いる	奴ら	は	正真正銘の
	Koko	*ni*	*iru*	*yatsu ra*	*wa*	*syousin syou mei no*
	This place LOCATIVE		be	guy PLURAL	TOPIC	authentic, real

(7)-b	厨房	ばっかり	だろう…
	chuubou	*bakkari*	*darou*
	middle school brats/kitchens	only	probably

Literal meaning: 'You guys who are here are probably all real kitchens….'

Channel 2 meaning: 'You guys who are here are probably all real immature middle school brats…' (From *Chuubou* Board)

3.1.3. Euphemism by *kanji* punning

This phonetic use of *kanji* shows users' deliberate avoidance of certain taboo and sensitive expressions. Before Channel 2 was created, there was a similar website, where certain expressions were automatically removed as *No-Good Words*. The example in Table 3, (a) *shine*, which is the imperative form of the verb *shinu* 'to die', was such a word. In order to avoid automatic deletion, users replaced the conventional character with another *kanji* having the same pronunciation.

Table 3. Formation process of euphemism by *kanji* punning

Taboo/ sensitive words	Euphemism	*Romaji*	Automatic *Hiragana* conversion	Deliberate *kanji* selection literal meaning	Channel 2 usage
a) 死ね *shin-e* 'die' imperative, 'drop dead'	氏ね *shi-ne* respect suffix + particle	*shine*	しね → *shine*	氏ね *shine* no such expression	氏ね *shine* 'drop dead, die'
	Variation: 詩ね	*shine*	しね →	詩ね 'poetry, isn't it'	詩ね 'drop dead, die'
	Variation: 士ね	*shine*	しね →	士ね no such expression	士ね 'drop dead, die'
	Variation: 市ね	*shine*	しね →	市ね 'city, isn't it'	市ね 'drop dead, die'
b) 金田一春彦 *kindaichi haruhiko* proper noun	禁打壱貼肥子 *kin-da-ichi-haru-hiko*	*kin da ichi haru hiko*	きんだいち はるひこ *kindaichi haruhiko*	禁打壱貼肥子 *kindaichi haruhiko* no such expression	禁打壱貼肥子 *kindaichi haruhiko* proper noun

Example (a) shows how *shine* 'die' is written. There are several different *kanji* for *shi*, because there are many characters that have the pronunciation [ʃi]. Therefore, users can deliberately choose any character that does not mean *death*. (8) shows an actual use of *shine* in a message from the 'Anything-goes' Board. In this case, the respect suffix *-shi* replaces the *kanji* for *die*.

Example (b) in Table 3 gives a phonetic use of *kanji* for a proper name. Since proper names can be found by computer search engines, different *kanji* with the same pronunciation are often used, in order to avoid discovery of the user's criticism or other negative comments about an individual or institution. In (b), taken from the Linguistics Board of Channel 2, we see the name of a well-known linguist, Haruhiko Kindaichi, written in completely incorrect *kanji* characters. This form of the name is used in (9), in which the author is making unfavorable comments about Kindaichi. By sight, it is very difficult to understand what the sequence of characters refers to, but if the pronunciation of each character is identified, it is not difficult to determine who or what is being discussed. In this way, Channel 2 has produced quite a few jargon uses of *kanji*, which can be understood easily by the in-group, but only with difficulty by outsiders.

(8) ネタ が フリー んだ よ。 氏 ね
neta ga huri: n da yo. shi ne
stuff SUBJECT old NOMINALIZER SFP RESPECT SUFFIX SFP
Literal meaning: '(I'm telling you) (your) stuff/information is old. *Non-word'
Channel 2 meaning: '(I'm telling you) (your) stuff/information is old/boring. Drop dead'
(From 'Anything-goes' Board)

(9) …禁打壱貼肥子 の 食べ物 日記 が 載っている。
…kindaichi haruhiko no tabemono nikki ga not-te iru
Kindaichi Haruhiko GENITIVE food diary SUBJECT appear-PROGRESSIVE
Literal meaning: Food diary of *kin-da-ichi-haru-hi-ko appears ….
Channel 2 meaning: Kindaichi Haruhiko's food diary appears ….
(From Linguistics Board)

3.2. Play with script shape similarity

We see the second type of unconventional usage on Channel 2, play with script shape similarity, in Table 4. Most of these examples of replacing one symbol with another involve the *katakana* syllabary, because some of the *katakana* symbols have very similar and sometimes confusing shapes.

A typical example is the replacement of *n* with *so*, shown in (a). This occurs because the two *katakana* symbols, *n* and *so*, look very similar; here Channel 2 users rely on the visual rather than phonetic information of scripts. The difference between *n* ン and *so* ソ is that in *n* ン, the longer slanting line goes upward from bottom to top, while in *so* ソ, the long line goes from the top down. In (a), we see the deliberate use of *so* for *n*. The original expression is *sumanai* 'sorry', and this gets shortened to *suman*. The last *katakana n* ン of *suman* looks similar to *so* ソ. Then, the users coin a new word, *sumaso*, relying on the shape similarity for other users' comprehension. *Sumaso* actually is a non-word in contexts other than Channel 2. Although there is no specific reason to prefer *sumaso* over *suman* as an expression of apology, the users' motivation, it seems, is to share the fun of creating and using visually-oriented play with words. The use of *sumaso* as an expression of apology was imitated by other users and has now been lexicalized as part of the common vocabulary in Channel 2. (b) and (c) are additional examples of this practice. An instance of *sumaso*, taken from the Movies Board, is given in (10). The writer of (10) has posted her message on the wrong thread, and so she apologizes, using *sumaso*.

Table 4. Formation process of coining new words based on shape similarity

	Origin	Shortening/ modification	Replacement due to shape similarity	Outcome literal meaning	Channel 2 usage
a)	スマナイ	スマン	ン → ソ	スマソ	スマソ
	sumanai	*suman*	*n* → *so*	*su ma so*	*sumaso*
	'sorry'	'sorry'		no such expression	'sorry'
b)	マジメ	マジ	ジ → ヅ	マヅ	マヅ
	mazime	*mazi*	*mazi* → *madu*	*madu*	*madu*
	'seriously'	'seriously'		no such expression	'seriously'
c)	スレッド	スレ	ス → ヌ	ヌレ	ヌレ
	sureddo	*sure*	*su* → *nu*	*nure*	*nure*
	'thread'	'thread'		no such expression	'thread'

(10) スレ　違い　スマソ　でした。
Sure *chigai* *sumaso* *deshita.*
Thread different/wrong sorry be. POLITE PAST

Literal meaning: Wrong thread, it was *sumaso
Channel 2 meaning: '(I'm) sorry (to have posted my message) in the wrong thread.'
(From Movies Board)

3.3. Conventionalized errors

3.3.1 Literacy errors

The third type of innovation on Channel 2 is found in conventionalized errors; at first they may have been used inadvertently, due to the user's insufficient literacy skills. Table 5 summarizes how one such error occurred. A Channel 2 user wrote *gaisyutsu*, instead of *kisyutsu*, which means 'already mentioned'. Apparently this user believed the *kanji* characters for the meaning 'already mentioned' were those in *gaisyutsu*, and so he or she entered the *romaji* for *gaisyutsu*. This was converted to *hiragana* and then the conversion system produced the *kanji* characters for 'going out of home/office', which are pronounced *gaisyutsu*. These characters were different from what the writer had in mind, so s/he decided to return to *hiragana*.

Table 5. Formation process for reading error

Intended *kanji*	Entered *Romaji*	Automatic *hiragana* conversion	Automatic *kanji* conversion literal meaning	Return to *hiragana*	Channel 2 usage
既出 *ki syutsu* 'above mentioned'	*gaisyutsu*	がいしゅつ *gaisyutsu*	外出 *gaisyutsu* 'going out of home'	がいしゅつ *gaisyutsu*	外出,がいしゅつ *gaisyutsu* 'above mentioned'

The actual message is given in (11), in which the writer uses *hiragana* for 'already mentioned'. After this message was posted, there was a tremendous reaction to it. (12) is the message immediately after (11). In (12), this very first responder asks whether the original sender meant *kisyutsu* 'already mentioned', and even adds the *hiragana* for *kisyutsu* 'already mentioned' in parentheses, to clarify the correct pronunciation, and to be sure that *kisyutsu* 'already mentioned' rather than *gaisyutsu* 'going out of home' was intended. The source of *gai* in *gaisyutsu* is either 概 or 慨, both of which are pronounced [gai] and share two of the same part or radical, 既 *ki*, in 既出 *kisyutsu.*

(11) がいしゅつ だった ら すみません。
Gaisyutsu *dat-ta* *ra* *sumimasen.*
Going out/already mentioned be-PAST CONDITIONAL apologize
Literal meaning: I'm sorry if (it [the following story]) has gone out of home.
Channel 2 meaning: I'm sorry if (it [the following story]) has already appeared.

(12)-1 がいしゅつ？
Gaisyutsu ?
going out
'(Are you saying) Going out?'

(12)-2 既出 (きしゅつ) の こと？
Kisyutsu [in *kanji*] (*kisyutsu*) [in *hiragana*] *no* *koto* ?
Already mentioned (already mentioned) of thing, fact
'Do you mean, "Already mentioned"?' (From Glossary of Channel 2 Terms, also archived in http://saki.2ch.net/news/kako/962/962871899.html)

The kind of interaction we see in (11) and (12) continued for a fairly long time, half a year after the message was first posted. The use of *gaisyutsu* 'going out' instead of *kisyutsu* 'already mentioned' was so amusing to many of the users that it stirred a considerable amount of discussion and interaction. Their reactions were, for example, "I can't stop laughing", "It's the birth of a new Channel 2 term", "I like *gaisyutsu* and so I'll use it," and so on. One other user praised the original user for his/her use of 'gaisyutsu', commenting on this use as a kind of pardon for those who might not have appropriate literacy. Based on the fact that this use has been inherited and favored by many others and has become conventionalized, we can infer that the standard forms may not always be the most welcome ones on this website. We can speculate further that users enjoy the freedom to create and use new words, unbound by standard convention.

3.3.1. Typing errors

The next and final example of conventionalized errors concerns typing errors. Table 6 shows how an unconventional way of writing 'me too' and another example came about.

Table 6. Formation process for typing errors

Intended *kanji*/expression	Entered *romaji*	Automatic *hiragana* conversion	Automatic *kanji* conversion literal meaning	Channel 2 usage	Lexicalized also as
a) 俺も 俺も *ore mo ore mo* 1st person pronoun + Particle 'me too me too'	*ore more mo*	おれもれも *Oremoremo*	俺 漏れ も *Ore more mo* 'I' 'leakage' 'too'	漏れも *more mo* 'me too'	漏れ *more* 'I, me' (without particle)
b) がんばる *ganbaru* 'to do one's best'	*gangaru*	がんがる *gangaru*	No corresponding *kanji* no such word	がんがる *gangaru* 'to do one's best'	がんがれ *gangare* 'do your best' (Imperative)

In Example (a) of Table 6, the first user intended to type *ore mo ore mo*, meaning 'me too me too'. When this user entered the words in *romaji*, s/he apparently did not enter all the roman letters, but skipped one, *o*, as shown in "Entered *romaji*". This *romaji* then got converted to the *hiragana* for *ore more mo*. The system correctly converted *ore* to the right character. But it took *more* to be the word for 'leakage'. Now *more* has been lexicalized to mean 'I' in Channel 2. The typing error in Example (b) above comes from typing *g* instead of *b*, because these two keys are near to one another on the keyboard. This error results in coinage of a new word that is in currency mostly in Channel 2. An example of how *more* instead of *ore* is actually used to mean 'I' is shown in (13). As we see, *more* is used even in serious discussion.

(13)	漏れ	は	(生成	スレ	に	も	少し	書いた …)
	more	*wa*	*(seisei*	*sure*	*ni*	*mo*	*sukoshi*	*kai-ta …)*
	leakage/I	TOPIC	generative	thread	LOC	also	a little	write-PAST

Literal meaning: As for leakage [I] (wrote a little also in the generative (grammar) thread) …
Channel 2 meaning: As for me, (I wrote a little also in the generative (grammar) thread) …
(From Linguistics Board)

4. Summary and conclusions

In this study I have analyzed the special linguistic and interactional features of online communication in one Japanese website. The unconventional uses of scripts on Channel 2 have at least two motivations: first, they are entertaining, and second, they are identity markers for this Internet subcultural community, an identity that can be characterized as free, funny, and unconventional. As we have seen, Channel 2 constitutes a community of practice, in which users are mutually engaged in constructing the website as their joint enterprise and have created a huge shared linguistic repertoire with unique uses of Japanese scripts. The kind of written informal, interpersonal, unedited communication in Channel 2, and on the Internet in general, allows us to explore new dimensions in linguistic and interactional behavior. This study is one step in this new direction for socio-cultural linguistics.

Primary sources

Channel 2 website: <http://www.2ch.net/2ch.html>
Glossary of Channel 2 Terms: <http://shake-hip.com/niwatori/>
Chuubou Board: < http://tmp.2ch.net/kitchen/>
Linguistics Board: < http://academy.2ch.net/gengo/>
Anything-goes Board: <http://tmp.2ch.net/mog2/>
PC News Board <http://pc.2ch.net/pcnews/>
Niten Purojekuto.[Channel 2 Dictionary Project Team] (2002). *Ni ten: ni channeru jiten. [Channel 2 Dictionary]*.Tokyo: Baacharu kurasutaa [Virtual Cluster] Publisher.

References

Nishimura, Yukiko. To appear. Linguistic innovations in informal online communication in Japanese. *Journal of Computer-Mediated Communication.*

Wenger, Etienne. 1998. *Communities of practice: learning, meaning, and identity.* Cambridge: Cambridge University Press.

Yukiko Nishimura
Toyo Gakuen University
1660 Hiregasaki, Nagareyama,
Chiba 270-0161 Japan

nisimura@hm.tyg.jp

Rethinking 'Thinking for Speaking'[1]

STÉPHANIE POURCEL
University of Durham, UK

0. Introduction

This paper makes use of Slobin's *Thinking for Speaking* theory (e.g. 1996a, 1996b, 1997, 2000, 2003b) as a stepping stone into the exploration of language and thought. If focuses on the linguistic realisation and on the cognitive conceptualisation of motion events. The present study suggests a different methodology of investigation, whereby in two separate experiments the stimuli and responses are cognitive only (Experiment 1) and linguistic as well as cognitive (Experiment 2). On the basis of results obtained with sixty-four native English speakers and seventy-five native French speakers, it questions the claims in Slobin's work regarding the differential salience of the dimensions of Path and Manner in the conceptualisation of motion events by speakers of different language types. Indeed the present results indicate that both English and French speakers showed an overall preference for Path in the required tasks.

1. Motion events

The morphosyntactic realisation of motion events differs across the world's languages in three main ways. These differences are caused by differential conflation patterns of the various dimensions of a motion in language. Talmy (1985) famously identified these dimensions as comprising of the Figure, the Ground, the Path, and the Manner of the motion. He further suggested a dual linguistic typology, whereby languages may be classed as being verb-framed if they typically conflate Path in the verb (e.g. Romance, Polynesian and Semitic languages), or satellite-framed if they conflate Path in a verb particle, or complement (e.g. non-Romance Indo-European languages).

Slobin (2003a) suggested a third equipollently-framed category corresponding to serial-verb languages, which conflate both Path and Manner in several verbs at once (e.g. Algonquian, Athabaskan, Hokan, Klamath-Takelman). Slobin (in

[1] This research has received approval from the Ethics Committee of the University of Durham (UK), and is funded by the Economic & Social Research Council (UK), award R42200154377.

press), however, later suggests considering 'a cline of Manner salience' and allocating languages along a continuum going from Path-salient languages to Manner-salient languages. Whichever theoretical approach may prove more convenient, the point here is that different languages demonstrate clear tendencies for favouring the allocation of one particular dimension of motion to one particular grammatical category, so that eventually, elements that are centrally expressed in the main verb in satellite-framed languages are found to be perfectly dispensable in verb-framed languages, for instance. Consider English and French in (1) and (2) below,

(1)	The bird	flew	out	of its cage.
	[Figure]	[Manner]	[Path]	[Ground]

(2)	L'oiseau	est sorti	de sa cage.
	[Figure]	[Path]	[Ground]
	The bird	*exited*	*its cage.*

In contrast to English, Romance languages are indeed what Slobin (in press) terms low-manner-salient languages, to the extent that these languages often do not express the concept of Manner of motion in language, unless it is relevant somehow to the semantic load.

2. Thinking for Speaking

Based on the above considerations regarding the domain of motion in linguistic expression, Slobin (e.g. 1996a, 1996b, 1997, 2000, 2003b) has put forward a neo-Whorfian reformulation of linguistic relativity. In this new approach, Slobin (1996a: 76) does not propose to examine "whatever effects grammar may or may not have outside of the act of speaking", i.e. non-linguistic thought; but, rather, he offers a new approach whereby

> the expression of experience in linguistic terms constitutes **thinking for speaking** – a special form of thought that is mobilised for communication. (ibid.)

His intention was thus to examine linguistic cognition, for which he gathered linguistic data from elicited narratives, natural discourse, creative fiction, translation work, and more, across a spectrum of subjects of different ages and linguistic backgrounds. These data confirmed Slobin's hypothesis that "in acquiring a native language, the child learns particular ways of thinking for speaking" (ibid.). Note however that no cognitive data is considered at any point.

Slobin's research illustrates how speakers of different languages are predisposed to attend to certain aspects of experience due to obligatory categories in grammar, but it does not address the question of the cognitive implications resulting from the use of particular languages. Rather, it points to correlates between grammar and thought. Yet even this much may be doubted since, as

Slobin points out, there is "nothing in the pictures themselves that leads English speakers to verbally express whether an event is in progress" (1996a: 88). But there definitely *is* something in the language they use that leads them to verbally express that an event is in progress, for instance. This something is precisely the lexicalisation patterns mentioned above, which speakers *must* follow in speech. This does not prove that speakers of different languages are more sensitive or pay more attention to certain aspects of the same reality than to other aspects (see Gennari et al. 2002: 55 for a similar evaluation).

In other words, Slobin's conclusions are circular from a psychological and relativistic perspective – even in the event that linguistic cognition is the object of investigation. The difficulty stems from his departing from a non-linguistic stimulus to arrive at a linguistic result – which was predictable – and from there to deduce (non-linguistic) patterns of habitual thinking. Overall, Slobin has demonstrated that speakers have to think about language itself in order to speak. This thinking becomes systematised to a certain degree in the process of language acquisition and use, and varies cross-linguistically according to specific grammars. In sum, Slobin showed *how* a specific language asks its users to highlight Path or Manner according to their native input. This does not by any means posit any cognitive consequences, and as such does not provide evidence for linguistic relativity. If the relativist's purpose is indeed to show that different languages and different fashions of speaking engender different ways of thinking, then their likely evidence ought to consist of those very ways of thinking (Lucy 1992). Observing linguistic behaviour merely helps document linguistic diversity, not cognitive divergences.

3. Motion conceptualisation

The question under debate is thus whether speakers of different languages conceptualise motion in terms of either Manner, or Path – or both – because of the restrictions imposed by their native language; or, in Slobin's words, do we obtain "divergent mental worlds for speakers of the two language types" (2000: 133), i.e. satellite- and verb-framed?

It seems essential to complement Slobin's extensive corpus data with some cognitive data speaking to the same issue. The present research therefore tested English and French native speakers on cognitive visualisation tasks. Two separate experiments were implemented. In the first one, subjects were asked to judge mute visual stimuli (video clips) in terms of similarity, e.g.

(3) a. a man running up a hill
    b. a man running down a hill
    c. a man walking up a hill

The test comprised fifteen such triads, with differing Manner and Path types, always organised so that two distinct pairs would correspond to Path similarity (e.g. 3a and 3c above) and to Manner similarity (e.g. 3a and 3b). Subjects were asked to identify the pair displaying more perceived similarity. In the second experiment, a second unrelated sample of subjects provided written linguistic descriptions of these stimuli prior to performing the similarity judgements.

It was expected that more Path associations would be found in the French sample of results overall, since Manner is a dispensable concept, linguistically. Whorfian relativism would also predict that this difference would be present in the first experiment. And, according to Thinking for Speaking, we should expect stressed differences in performance between experimental formats, since Experiment 2 involves explicit language.

4. Experimental results

4.1. Item analysis

All the results display a clear preference for Path associations (see Table 1).[2]

Table 1. Proportions of association types in Experiments 1 and 2.

		Manner associations	Path associations
Experiment 1	English (N=34)	37%	59%
	French (N=35)	37%	56%
Experiment 2	English (N=30)	44%	52%
	French (N=40)	42%	51%

From the above, it appears that results do not differ according to the native language of the subjects, but rather according to the nature of the task. Indeed, the preference for Path associations is less marked in the second experiment, that is, when linguistic descriptions are required prior to the similarity judgements. This may reflect an influence of explicit language on performance, as Thinking for Speaking would have predicted, though it is not quite clear how, as the direction of the data is the same for the two language groups.

Another possible explanation for the present state of affairs is that subjects in Test 2 were presented with a more complex task than subjects in Test 1; they may therefore have paid attention to more elements in the motion scenes, or they may simply have been distracted by the extra task. In brief, the more controlled and the more complex the task, the less spontaneous and realistic the answers. It may be then that the present differences reflect task complexity issues rather than anything else.

4.2. Subject analysis

Subjects displayed dominance patterns in their association choices concordant with the item analysis (see Table 2).[3]

[2] Note however that these differences are not statistically significant. The discussion therefore focuses on preferential patterns.

[3] Dominance patterns were established on the basis of at least two differential associations, e.g. 6 Manner associations and 9 Path associations would indicate a Path dominance.

Table 2. Dominance patterns in Groups 1 and 2.

		Manner dominance	Path dominance	No dominance
Group 1	English (N=34)	12%	67%	21%
	French (N=35)	14%	66%	20%
Group 2	English (N=30)	33%	47%	20%
	French (N=40)	25%	50%	25%

Again, the figures indicate a striking correlation between responses and experimental format, rather than between responses and linguistic motion typologies.

Individuals' knowledge of languages other than their native tongue was also monitored for potential linguistic effects. Both language groups provided monolingual speakers, bilingual speakers, and speakers with non-fluent knowledge of other languages.[4] There appeared however to be no consistent correlation between individuals' performance and their knowledge of languages of the opposite typological category. The nature of the task remains therefore the obvious factor responsible for the divergence in the above dominance patterns.

4.3. Linguistic analysis

Linguistically, both the French and the English Group 2 confirmed the typological framework drawn by Talmy, and the findings made by Slobin (see Table 3).

Table 3. Motion elements conflated in the verbs used in Group 2's descriptions.

	Manner in verb	Path in verb
English Group 2	85%	15%
French Group 2	33%	65%

These figures offer support to Slobin's idea of a cline of dimensional salience, as neither language groups conforms to Talmy's categories in the absolute. Indeed, it is perfectly acceptable to conflate Path in the English verb without any mention of Manner in the remaining of the sentence, e.g.

(4) A man opens the door.

Likewise, French allows for Manner verb constructions without Path cues, e.g.

(5) Un homme court.
A man is running.

In fact, French appears to be rather flexible in its treatment of motion expression. The figures in Table 3 are indicative of this, as only 65% of the descriptions provided seem to conform to typological predictions. It is important nonetheless

[4] Note that only the languages belonging to the opposite typological pattern were of interest here.

to recall that the present data was elicited in highly controlled lab conditions, and are not reflective of natural discourse practices (see Slobin in press for a more thorough investigation of natural data). However the present data illustrates the linguistic patterns available to native speakers for expressing motion, and these are more numerous and complex than Talmy's clean typology seems to suggest. Not only can French speakers conflate Manner in the verb without elaboration on Path information, as in (5), but they can also produce verb-framed reverse patterns, whereby Manner is conflated in the verb and Path is encoded in an adjunct, such as a gerund, e.g.

(6) Un homme court en traversant la rue.
A man runs crossing the road.
A man is running across the road.

This pattern was reproduced in several instances, and often even generated odd constructions, e.g.

(7) Un homme pédale à vélo en montant.
A man pedals on his bike going up.
A man is cycling up (the road).

These sentences may be a mere reflection of the experimental conditions, though their formal format would have predicted otherwise. They may also suggest a process of latent change as has been witnessed with Latin for instance (Slobin 2003a). Or they may simply indicate a greater flexibility of motion expression than had originally been attributed to French. Overall, these results may be suggestive of a hybrid system, whereby French displays characteristics of both verb- and satellite-framed patterns. Kopecka (2003) also exemplified this possibility with the existence of Path verbal prefixes in French, e.g. *déshabiller (undress), survoler (fly over), revenir (come back), soulever (lift up),* etc. – though it is unclear as yet whether these prefixes may be considered satellites.

For the present purposes, the main interest in examining these descriptions is to assess the parallels between categorisation choices and linguistic categories (see Table 4).

Table 4. Proportions of matches between linguistic elements and associations.

	Verbs	Satellites	Optional constitu-ents	Different constitu-ents	Noun phrases	No parallel
English	38%	46%	8%	1%	0%	7%
French	42%	N/A	28%	6%	1%	26%

An interesting finding here is that association types did not typically parallel similarity in the main verb. In fact, associations fail to match verbal similarity in

62% of cases in the English data and in 58% of cases in the French data. Similar findings have been reported by Papafragou et al. (2002: 210) who noted that "[i]n neither language groups did verb matches correlate with categorisation performance." It is also interesting to note that in a rather substantial number of instances, there was no correlation at all between the linguistic information provided by the subjects and their association choices.

This lack of consistent correlated patterns suggests that language does not have a pervasive influence over associative thinking. It has been suggested that subjects tend to appeal to linguistic categories when prompted to use language in categorisation tasks using non-linguistic stimuli, such as in Experiment 2 (Papafragou et al. 2002: 216), so that language becomes an active tool in problem-solving tasks. Gennari et al. (2002) draw similar conclusions and report findings of linguistic effects in similarity tasks after linguistic encoding concordant with Talmy's typology. The present findings, however, fail to obtain these linguistic effects. Some effects are present but they do not match with the typology, which suggests that language is not the factor biasing performance towards fewer Path associations in Experiment 2. Furthermore, the present experiments show that the use of language as a strategic resource to solve associative tasks is not systematically employed by subjects. Debriefing subjects after the tests confirmed this and indicated a wide variability in individual cognitive approaches.

Finally, the linguistic data was analysed in the light of Talmy's definition of semantic salience (1985: 122), whereby

> a semantic element is backgrounded by expression in the main verb or in any closed-class element (including a satellite – hence anywhere in the verb complex). Elsewhere it is foregrounded.

As Table 4 suggests, subjects often 'foregrounded' information relating to motion in optional constituents, e.g.

(8) A man is walking *with a limp.*

According to the above principle, we may expect parallels between the type of foregrounded information and the nature of the corresponding association (see Table 5).

Table 5. Foregrounding and association types in the English data.

	Manner foregrounded	Path foregrounded
Manner association	29	20
Path association	16	9

Overall, Talmy's principle would predict that foregrounded Path information would bias judgements towards a Path association choice. However, as Table 5 clearly illustrates, this pattern does not obtain. Talmy's principle of salience represents a linguistic understanding of salience but by no means does linguistic salience entail cognitive salience, as the present results illustrate.

4.4. Aspect

As highlighted by Aske (1989), among others, there exists two distinct types of Path, namely telic and atelic (or locative) Paths. Telicity entails an end-point or the crossing of a boundary, so that (9) represents a telic type of Path, and (10) an atelic one:

(9) We walked into the room.

(10) We walked along the beach.

In the present stimuli, four triads depicted atelic types of Path, eight triads depicted telic Paths, and another three triads combined the two types. Both language groups displayed a clear correlation between the type of Path and the type of association made, whereby telic stimuli show Path preferences and atelic stimuli show Manner preferences (see Table 6). Conceptually, the fact that telicity entails reaching an end-point may explain the logic behind the relatively greater salience of Path/ Result. And conversely, in atelic scenes, no end-point being reached, the process or Manner may naturally seem a more important dimension of the motion. Another possible explanation resides in the neural structures of the brain (i.e. pre-motor cortex), as it is indeed the same structures that are involved in the processing of complex motion and of aspect in language (Lakoff 2002).

Table 6. Proportion of association types relative to telicity.

		Telic sets		Atelic sets	
		Path	Manner	Path	Manner
Test 1	English (N=34)	67%	33%	43%	57%
	French (N=35)	64%	36%	43%	57%
Test 2	English (N=30)	56%	44%	34%	66%
	French (N=40)	55%	45%	33%	67%

All four experimental groups showed that dimensional salience is relative to aspect. We may therefore infer that there is a definite link between the conceptual salience of either Manner or Path and telicity in motion. This suggests that neither Manner nor Path is more conceptually salient of its own. Rather, the perceived salience of either dimension is dependent on aspectual features of the motion event in question. This also undermines the absolute acceptability of Path as the core schema of motion (see Talmy 1991). Indeed, though Path seems critical in conveying motion semantics, it is not indispensable as such for the expression and the conceptualisation of motion, as sentences such as (5) above illustrate.

5. Conclusion

The findings presented in this paper challenge some of Slobin's early relativistic postulations concerning the relatively higher cognitive salience of Path over Manner to verb-framed language speakers (e.g. 1996: 88, 2000: 133). They show that both English and French speakers favour Path as the most important conceptual element in motion events. This runs counter to the linguistic emphasis on Manner in English and on Path in French, and thus suggests that basic motion conceptualisation is the same for speakers of both language categories. In other words, language does not appear to play any significant role in the perceived salience of dimensional features of motion.

On the other hand, it is important to see that the motion scenes used in the present stimuli contrasted 'absolute', or non-discrete, Manner and Path types, e.g. *walking* vs. *running, down* vs. *up*. It is conceivable that linguistic effects may be found when contrasting discrete, or 'fine-grained', Path and Manner types, e.g. different types of *walking* based on force dynamics, speed and the like (see Slobin 2003c). Oh (in progress) is currently obtaining such results in recall tasks on fine-grained Manner distinctions, in which Korean speakers show higher error rates than English speakers. This area may indeed prove more promising for Whorfian effects. The current research therefore presents preliminary results, and by no means pretends to dismiss a relativistic reality.

References

Aske, Jon. 1989. Path predicates in English and Spanish: A closer look. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*, 1-14.

Gennari, Silvia, Steven Sloman, Barbara Malt & W. Tecumseh Fitch. 2002. Motion events in language and cognition. *Cognition* 83: 49-79.

Kopecka, Anetta. 2003. The semantic structure of motion verbs in French: typological perspectives. Conference presentation, Space in Languages: Linguistic Systems and Cognitive Categories, Paris, 08.02.03.

Lakoff, George. 2002. The cog hypothesis. Conference presentation, Sixth Conceptual Structures, Discourse & Language Conference, Houston TX, 13.10.02.

Lucy, John. 1992. *Language Diversity and Thought*. Cambridge: University Press.

Oh, Kyung-ju. In progress. Language, cognition, and development: Motion events in English and Korean. PhD dissertation, UC Berkeley.

Papafragou, Anna, Christine Massey & Lila Gleitman. 2002. Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition* 84: 189-219.

Slobin, Dan. 1996a. From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (eds.) *Rethinking Linguistic Relativity*, 70-96. Cambridge: University Press.

Slobin, Dan. 1996b. Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S.A. Thompson (eds.) *Grammatical Constructions: Their Form and Meaning*, 195-217. Oxford: University Press.

Slobin, Dan. 1997. Mind, code, and text. In J. Bybee, J. Haiman & S.A. Thompson (eds.) *Essays on Language Function and Language Type: Dedicated to T. Givón*, 437-67. Amsterdam: John Benjamins.

Slobin, Dan. 2000. Verbalised events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (eds.) *Evidence for Linguistic Relativity*, 107-38. Amsterdam: John Benjamins.

Slobin, Dan. 2003a. What makes manner of motion salient? Conference presentation, Space in Languages: Linguistic Systems and Cognitive Categories, Paris, 08.02.03.

Slobin, Dan. 2003b. Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (eds.) *Advances in the Investigation of Language and Thought*. Cambridge, MA: MIT Press.

Slobin, Dan. 2003c. From infancy to adulthood: Exploring the effect of linguistic input on the discrimination of Manner and Path in motion events. Symposium presentation, Biennial meeting of the Society for Research in Child Development, Tampa, FL, 26.04.03.

Slobin, Dan. In press. The many ways to search for a frog: linguistic typology and the expression of motion events. In S. Strömqvist & L. Verhoeven (eds.) *Relating Events in Narrative: Typological & Contextual Perspectives*. Mahwah, NJ: LEA Publishers.

Talmy, Leonard. 1985. Lexicalisation patterns: Semantic structure in lexical forms. In T. Shopen (ed.) *Language Typology and Syntactic Description*, vol. 3, 57-149. Cambridge: University Press.

Talmy, Leonard. 1991. Path to realisation: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 480-520.

School of Linguistics and Language
University of Durham
Elvet Riverside
Durham, DH1 3JT
United Kingdom

s.s.pourcel@durham.ac.uk

The Japanese 'Nominative' Particle: A New Approach

BENJAMIN M. ROSENTHAL
University of California, Los Angeles

0. Introduction

The Japanese case particle *ga* tends to be glossed in the literature as either NOMINATIVE or SUBJECT. Although the application of such well-established labels to Japanese case particles facilitates the drawing of parallels between Japanese and other languages, it does not give us a clear picture of how particles like *ga* actually function.

At least as far back as Kuno 1973, the prevailing view has been that *ga* indicates subject, but also the object of certain predicates that refer to preferences (e.g. *suki* 'like'), desires (e.g. verbs with the affix *-ta-* 'want'), needs (e.g. *iri-* 'need'), or abilities (e.g. verbs with the affix *-e-* 'can'). Thus, in theory, a single predicate can have two participants marked by *ga*. The use of the term 'nominative' allows one to gloss *ga* consistently, whether the case relationship is subject or object (Sugamoto 1982). However, one gets the impression that the object-marking function of *ga* is seen as being a quirky anomaly, as when Shibatani (1990: 305) says "the normal correspondence between the nominative *ga* and the syntactic subject is disrupted in a number of construction types."

There have been at least two major challenges to the prevailing view of *ga*: Aoyama 1982 and Ono et al. 2000. Both analyses eschew the view that *ga* is a 'case particle', i.e. a particle that contributes information on the relationship of a participant of a predicate to that predicate. However, they fail to take into account how *ga* contrasts with other case particles such as *o* and *ni*.

In order to address the issues raised by Ono et al., whose analysis is based on actual data from conversation (rightly regarded as the most basic genre of language), this paper looks strictly at conversation. Evidence is adduced from three corpora of native adult-adult Japanese conversation.[1] This paper takes a "conservative" approach in that it sees *ga* as a case particle and that it therefore needs to be understood in relation to other case particles. However, it takes the

[1] JPN (University of Arizona), Pac Rim (University of California, Santa Barbara) and Callfriend (downloaded from www.talkbank.org).

radical view that *ga* does not signify subject per se, but is the case particle used with subject nouns because other case particles would specify 'non-subject'.[2] This paper is also novel in that it takes an integrated view of *ga* as being conditioned by two sets of factors: the participant-predicate relationship and the pragmatic environment.

1. Previous Analyses

1.1. Aoyama (1982)

According to Aoyama, *ga* is properly defined in opposition to the particle *wa*. He characterizes this opposition in terms of 'focus', or concentration of attention, on the marked item: *wa* signifies HIGH FOCUS, while *ga* signifies LOW FOCUS.

However, Aoyama overlooks some crucial facts. The particle *ga* says something about the relationship of a noun to a predicate, and is thus constrained in terms of which nouns it can mark:

(1) (talking about why a subject in school has become more enjoyable)
uchira ga tabun nareta n da to omou yo
we GA probably get.used.to-P NZ CO-NP Q think-NP PP [3]
"I think it's that we've gotten used to it."
[JPN]

(2) (talking about some photographs)
dare ni miseta no
who NI show-P NZ
"Who did you show them to?"
[Pac Rim]

In (1), *ga* clarifies that the noun *uchira* 'we' refers to the entity that gets used to something. If, for example, the particle *ni* were used in place of *ga*, it would signify that 'we' is the entity that is gotten used to. Conversely, in (2), *ni* clarifies that 'who' refers to the identity of those to whom some photographs were shown; if *ga* were used instead, it would signify that 'who' refers to the identity of the one who showed the photographs. Thus, we can say that for the verbs *nare-* 'get used to' and *mise-* 'show', *ga* and *ni* are 'designated' for different roles.[4]

The particle *wa*, meanwhile, can mark a noun in a role for which *ga* is designated:

[2] I find that in Japanese, there is no need to draw a distinction between the categories 'noun' and 'noun phrase' (NP).
[3] The following glosses are employed in this paper: CO = copula; EP = epistemic auxiliary; LK = linking particle; NG = negative morpheme; NP = non-past morpheme; NZ = nominalizer; P = past morpheme; PP = pragmatic particle; PS = passive morpheme; Q = quotative particle.
[4] The designation of case particles is sensitive not only to the lexical core of the predicate, but also to certain inflectional affixes such as the 'passive' affix *-are-*. This issue will be addressed in §3.

(3) *Shibuya no koosaten wa itsumo-no-yoo-ni konde-ita no*
Shibuya LK intersection WA as.always get.crowded-be-P NZ
"The intersections in Shibuya were crowded as always."
[Pac Rim]

For the verb *komi-* 'get crowded', *ga* is designated for the role of 'that which gets crowded'. However, *wa* can also occur with nouns in roles for which other particles are designated:

(4) *tooingu wa inshuransukampanii ga tanonderu kara moo*
towing WA insurance.compnay GA request-be-NP so already
"As for towing, the insurance company has already requested it, so (that's been taken care of)."
[JPN]

The *wa*-marked noun, *tooingu* 'towing', is in the role of 'that which is requested', for which the designated particle is not *ga* but *o*. However, *ga* is designated for the role of 'that which requests', and it occurs in that capacity in (4). In neither (3) nor (4) does the presence of *wa* clarify the relationship of the marked noun to the predicate; that relationship is ascertained entirely through inference. In fact, case particles other than *ga* and *o* can actually co-occur with *wa*:

(5) *Akira-kun ni wa hanashitenai daroo ne*
Akira NI WA speak-be-NG-NP EP PP
"I probably haven't told you, Akira."
[Pac Rim]

For the verb *hanashi-* 'speak', *ni* is the designated particle for "that which is spoken to". It is therefore *ni*, not *wa*, that clarifies that 'Akira' is the entity to which something has not been spoken. Thus, it would be difficult to defend the idea that *wa* as well as *ga* directly contributes information on noun-participant relationships.

The distribution of *wa* can be explained in terms of scope: it limits the scope of a predicate to the marked item, thereby making that item 'stand out'. For example, the speaker of (3) is describing a scene that took place during a holiday period in Shibuya, a trendy area of Tokyo. Just prior to this utterance, she had said that most Tokyo residents were out of town, but a lot of people were in Shibuya from the countryside. The use of *wa* highlights the intersections of Shibuya as being crowded, because during this holiday period one would expect that most intersections in Tokyo would not be crowded. Thus, it is clear that *ga* should be understood in opposition not to *wa* but to other particles like *o* and *ni*.

1.2. Ono et al. 2000

Ono et al. investigated the occurrence of *ga* in spontaneous conversation among adult native speakers of Japanese and made the following findings: (a) more than two thirds of the time, *ga* does not appear with a noun in a role for which *ga* is a designated case particle; (b) *ga* occurs mainly with subjects of intransitive verbs, rarely occurring with subjects of transitive verbs; and (c) *ga* occurs chiefly with non-verbal predicates (i.e. adjectives and nouns), motion verbs (e.g. *ki-* 'come') and verbs used to establish the existence of a referent (e.g. *ari-* 'be/exist', *deki-* 'come into being'). Furthermore, they found that the tokens of *ga* in their data could be attributed to pragmatic motivations, such as: (a) potential ambiguity of reference; (b) the introduction of a new referent; (c) a morphologically complex noun (e.g. a nominalized clause); and (d) the singling out of the referent from among other potential referents (what Kuno 1973 identifies as 'exhaustive listing'). Based on these findings, Ono et al. conclude that *ga* does indeed mark a noun as a participant of a predicate, but only in "pragmatically highly marked" environments, with ∅ (no particle) being the unmarked choice, and that *ga* therefore "has very little to do with either case or grammatical relations" (p.78).

Although their analysis reveals much about the use of *ga* in conversation, Ono et al. do not fully account for its use in that genre because they do not contrast it with other case particles. Specifically, they do not ask the question, "when *ga* does occur, why is it *ga* and not some other case particle like like *o* or *ni*?" In most of their examples, such as (6), the substitution of *ga* with another case particle would either result in the indication of a different noun-predicate relationship or be incoherent:

(6) (talking about some photographs)
ṇanka nikkori-to suggoi kawaiku waratten no ga atta n da yo ne
like wide-grin really cutely grin-be-NP NZ GA be-P NZ CO-NP PP PP
"There was one of you with like a really cute, wide grin, remember?"
[Ono et al.'s (11); gloss and translation modified from original]

Example (6) illustrates some of the properties of *ga* identified by Ono et al., i.e. marking a complex noun that is the subject of an intransitive verb used to establish the existence of a referent. However, any other case particle but *ga* would signal that the marked noun was not the subject of the predicate *ari-* 'be'. Thus, while *ga* may only appear in certain kinds of environment in conversation, it is not sufficient to say that it marks a noun simply as "a" participant of a predicate; it must mark the noun as "a certain kind of" participant. It would therefore be premature to dismiss the relevance of "case relations" to *ga*.

2. New Approach to *ga*

Ono et al. (2000) acknowledge that the attested skewing of *ga* towards intransitive rather than transitive subjects is in part attributable to pragmatic factors that favor a higher occurrence of intransitive subjects in Japanese. In a separate study of two-party face-to-face conversations in Japanese between young adult native speakers, Matsumoto (1997) found 764 tokens of intransitive predicates (intransitive verbs and non-verbal predicates), compared to only 357 tokens of transitive verbs. Moreover, while 320 (42%) of the intransitive predicates had overt subjects, only 48 (13%) of the transitive verbs had overt subjects. While Matsumoto does not give figures for how many of those transitive and intransitive subjects were marked with *ga*, it is clear that in her data, there were many more opportunities for *ga* to appear with an intransitive than a transitive subject. Matsumoto's study thus suggests two factors contributing to the skewing of *ga* towards intransitive subjects in conversation: a tendency for intransitive predicates to outnumber transitive predicates, and a tendency for intransitive subjects to be mentioned more often than transitive subjects.[5]

Clearly, then, the attested skewing of *ga* towards intransitive subjects should not be taken as evidence that there is no meaningful correlation between *ga* and the notion 'subject'. This is made even more apparent by the fact that in the kind of data analyzed by Ono et al., one **can** find tokens of *ga* marking transitive subjects. For example, the token of *ga* in (4) marks the subject of the transitive verb 'request', and the token in (7) below marks the subject of the verb *tatekae-* 'front (payment)':

(7) (talking about how a friend had to cancel a trip)
kyanseru-ryoo no goman-en wa kareshi ga tatekaeta
cancellation-fee LK 50,000-yen WA boyfriend GA front-P
"As for the cancellation fee, her boyfriend fronted it."
[JPN]

Thus, *ga* is compatible with transitive as well as intransitive subjects. What Ono et al. have discovered then is a **tendency**, not a hard and fast rule.

If we consider Ono et al.'s findings together with the evidence presented here, what begins to emerge is a picture of *ga* in which its appearance in conversation is conditioned by two sets of factors: the participant-predicate relationship on the one hand, and the pragmatic environment on the other. The participant-predicate relationship determines the choice of *ga* over other case particles such as *o* and *ni*, while the pragmatic environment determines the choice of *ga* over a non-case particle like *wa* or simply Ø (no particle). This picture of *ga* is supported by evidence that a similar picture can be drawn of at least some of the other case particles. Fujii and Ono (2000) analyzed the occurrence/non-occurrence of the case particle *o* (commonly glossed as ACCUSATIVE or DIRECT OBJECT) in conversation and found that *o* too occurs in less than one third of all places where it theoretically could, and again that its occurrence tends to be motivated by a

[5] For further discussion of the first tendency, see Thompson and Hopper 2000; for further discussion of the second tendency, see Matsumoto 1997.

limited set of pragmatic motivations (singling out a referent from among other potential referents and facilitating cognitive processing). Likewise, the case particle *ni* (commonly glossed as DATIVE or INDIRECT OBJECT) does not always occur where it theoretically could in conversation:

(8) *moo Nihon Ø kaettara gyaku-karuchaashokku de sugoi*
already Japan return-when reverse-culture.shock CO incredible-NP
"When we go back to Japan, the reverse culture shock will be incredible."
[JPN]

The particle *ni* is a designated particle for the 'goal/destination' role of *kaeri-* 'return'. Lee (2002) recognizes that *ga*, *o*, and *ni* do not always occur in conversation where they theoretically could, and regards the absence of any one of the three particles as a single phenomenon (a 'zero particle'). Thus, occurrence/non-occurrence in accordance with pragmatic factors is not unique to *ga*, and is a separate issue from what *ga* signifies regarding a participant-predicate relationship.

Now let us turn to the issue of what *ga* says about a participant-predicate relationship. If neither Aoyama 1982 nor Ono et al. 2000 offer an alternative explanation, then we are still left with the proposition that there is a "normal correspondence" between *ga* and the participant-predicate relationship 'subject'.

Clearly, a one-to-one correspondence cannot be established; otherwise, it should not be possible for *ga* to be a designated case particle for more than one participant role. However, if we take the notion of 'subject' rather than *ga* as our starting point, we find that there is indeed a correlation: if a noun is the subject of a predicate, then the only compatible case particle will be *ga*. We have already noted that *ga* is the only case particle compatible with the subject of (6). The same is true for (1), (4), and (7), where the *ga*-marked noun is also the subject of the predicate. Furthermore, if the verb is passivized, in which subject status is shifted away from the default participant role, then *ga* becomes the only viable case particle for the passive subject:

(9) (talking about a friend's chances of getting into medical school)
<u>*waku*</u> <u>*ga*</u> *kime-<u>rare</u>-te-te sono waku ga mata chiisai n da tte*
<u>quota</u> <u>GA</u> decide-<u>PS</u>-be-and that quota GA again small-NP NZ CO-NP Q
"Apparently, there's a fixed quota, and that quota is small."
[JPN]

Were it not for the presence of the passive morpheme *-rare-*, the first token of *ga* in (9) would be incongruous. This is because the role occupied by *waku* with respect to the verb *kime-* 'decide' (i.e. 'that which gets decided') is not the default subject role for that verb; as such, the designated case particle would be *o* rather than *ga*. Therefore, if a participant role is the subject, whether by default or derivation, the only compatible case particle will be *ga*.

To say that *ga* is the only case particle designated for 'subject' does not entail that *ga* is restricted to subject nouns. All it entails is that other case particles are incompatible with 'subject'. This, then, appears to be the crucial distinction between *ga* and other case particles: other case particles specify a non-subject relation. This difference can be characterized as one of **semantic specificity**: *ga* is the only case particle not to specify that the marked noun is not the subject of the relevant predicate. This accounts for the fact that *ga* is not restricted to subject nouns; *ga* does not specify 'subject', it only refrains from specifying 'not subject'.

None of this is to say that the subject role **must** always be marked by the designated case particle. For example, if (1) were uttered without any particle on 'we', then it would be up to the hearer to figure out the relationship of 'we' to the verb.

(1') *uchira Ø tabun nareta n da to omou yo*
we probably get.used.to-P NZ CO-NP Q think-NP PP
"I think it's that we've gotten used to it."
[JPN]

Given the context, the logical inference would be that 'we' are the ones who have done the 'getting used to'. The presence of *ga* merely **eases** the burden of inference. In this way, the relative infrequency of *ga* in conversation is not inconsistent with its signification of a participant-predicate relationship.

3. Conclusion

This paper takes a fresh approach to the analysis of the Japanese nominative particle *ga* in two ways. First, it presents an integrated view of *ga* as being conditioned by two sets of factors: the participant-predicate relationship and the pragmatic environment. The choice of *ga* over other case particles in conversation is a reflection of the participant-predicate relationship, while the actual appearance/non-appearance of *ga* is a reflection of the pragmatic environment (as revealed by Ono et al. 2000). Second, this paper clarifies the relationship of *ga* to the notion 'subject': if a participant role of a predicate is also the subject role, either by default or through derivative morphology, then *ga* is the designated case particle for that role. The fact that *ga* is not limited to subject nouns can thus be explained in terms of semantic specificity: it is not that *ga* specifies a subject relation, but rather that *o*, *ni* and other case particles specify non-subject relations, leaving *ga* as the only option for subject nouns.

It is reasonable to hypothesize that if *ga* is found to have different rates of occurrence in genres other than conversation, it will be attributable to a different set of pragmatic factors (e.g. conversation affords access to interactional resources like prosody and body language that are not readily available in written discourse). However, that is a matter for future research.

References

Aoyama, Takashi. 1982. The system of focus in Japanese. *Columbia Working Papers in Linguistics* 7: 91-106.

Fujii, Noriko and Tsuyoshi Ono. 2000. The occurrence and non-occurrence of the Japanese direct object marker *o* in conversation. *Studies in Language* 24: 1-39.

Kuno, Susumu. 1973. *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

Lee, Duck-Young. 2002. The function of the zero particle with special reference to spoken Japanese. *Journal of Pragmatics* 34: 645-682.

Matsumoto, Kazuko. 1997. NPs in Japanese Conversation. *Pragmatics* 7: 163-181.

Ono, Tsuyoshi, Sandra A. Thompson, and Ryoko Suzuki. 2000. The pragmatic nature of the so-called subject marker *ga* in Japanese: Evidence from conversation. *Discourse Studies* 2(1): 55-84.

Shibatani, M. 1990. *The Languages of Japan.* Cambridge: Cambridge University Press.

Sugamoto, N. 1982. Transitivity and Objecthood in Japanese. In P.J. Hopper & S.A. Thompson (eds.) *Studies in Transitivity*, 423-447. New York: Academic Press.

Thompson, Sandra A. and Paul J. Hopper. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. In J. Bybee & P. Hopper (ed.) *Frequency and the Emergence of Linguistic Structure*, 27-60. Philadelphia: John Benjamins.

Benjamin M. Rosenthal
Department of East Asian Languages and Cultures
290 Royce Hall
University of California, Los Angeles
Los Angeles, CA 90095-1540

milaben@hotmail.com

Lexical Confusability and Degree of Coarticulation

REBECCA ANNE SCARBOROUGH
University of California, Los Angeles

1. Listener-motivated accommodations in speech

Speech is inherently communicative—"we speak to be heard in order to be understood" (Jakobson, Fant, & Halle, 1952). Thus, it stands to reason that speech production should be influenced by a speaker's desire to be understood. A speaker must produce a signal from which his listener can recover the intended message. If he is not sufficiently careful, communication will be unsuccessful. On the other hand, as long as a speaker remains intelligible to his listener, there is no reason that he cannot adjust his production to reduce the amount of effort he must expend as speaker. Lindblom (1990) has proposed a model of the interaction among the forces that shape speech, characterizing speech communication as a dynamic balance between speaker-oriented and listener-oriented forces, regulated by the communicative context. In other words, as factors in the communicative situation place extra demands on the listener, decreasing his chances of recovering the message, the speaker must adjust his pronunciation in order to produce clearer speech (referred to in Lindblom's model as "hyper-speech"). However, when conditions are favorable for communication, the speaker is free to conserve articulatory effort, producing reduced speech (or "hypo-speech").

Research has shown that speakers are sensitive to a number of different types of listener difficulties and make corresponding acoustic-phonetic accommodations. One early and easily confirmable observation of a listener-motivated speech accommodation was that people talk more loudly (Lombard 1911) and more slowly (Hanley & Steer 1949) in noisy environments than in quiet ones. Similarly, speech directed toward hearing-impaired listeners is slower and less phonologically reduced than normal conversational speech (Picheny, et al. 1986).

Importantly, these ostensibly listener-motivated accommodations (or at least the speech containing these accommodations) have also been shown to have an observable positive effect on intelligibility for listeners. Experiments have verified, for example, that speech produced in noise is more intelligible when presented at a constant speech-to-noise ratio than speech produced in quiet (e.g., Summers, et al. 1988). And Picheny, et al. (1985) present evidence of a substantial improvement in intelligibility for clear speech relative to normal conversational speech for hearing impaired listeners.

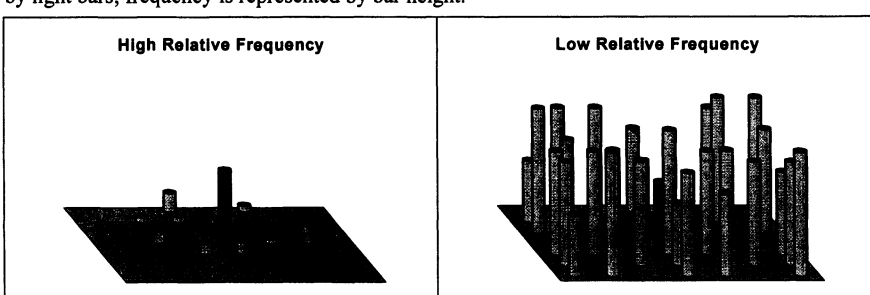
Factors internal to the structure of an interaction may also motivate speaker accommodations. Words that are less predictable from the conversational context are more intelligible when removed from their context and presented in isolation than are more predictable words (Lieberman 1963), and the first occurrence of a word in a narrative is more intelligible than the second occurrence of the same word when the words are heard out of their discourse context (Fowler & Housum 1987). These results have been interpreted as suggesting that new information and unpredictable information are spoken more clearly than old or predictable information because speakers know that listeners will be unable to rely on inferences from context to provide top-down cues to help them in lexical perception.

2. Lexical confusability

Factors at an even "lower", lexical level may also increase or decrease the likelihood of a word being correctly identified by a listener. For instance, more frequently occurring words are more quickly and accurately understood than less common ones (e.g., Gordon 1983; Glanzer & Eisenreich 1979). Intelligibility is further mediated by the number of words that are phonologically similar to a given word (i.e., the number of neighbors): the more neighbors a word has, the harder it is for a listener to identify (e.g., Luce 1986; Pisoni, et al. 1985). These two kinds of effects on the speaker-independent intelligibility of a word can both be understood as arising from the competition intrinsic in the process of lexical access.

Word identification involves discriminating among various entries in the mental lexicon. The difficulty of accessing a particular lexical item is related to the probability of correctly recognizing that item from among its lexical neighbors. Models of lexical access such as the Neighborhood Activation Model (Luce 1986; Luce & Pisoni 1998) propose that the frequency of a word and the number and frequencies of its competitor neighbors have various excitatory and inhibitory effects in lexical access that affect the activation of a target word relative to its competitors. The combined effects of frequency and neighborhood on the relative activation of a word can be represented by a ratio of the frequency of the word to the sum of the frequencies of all its neighbors (referred to here as relative frequency or *R*). The figure in (1) below illustrates high and low relative frequencies. In the high-*R* picture, the target word (which has a higher frequency than its few and infrequent neighbors) stands out, while in the low-*R* picture, the target word (with its relatively low frequency and many frequent neighbors) is obscured by its neighbors. With respect to intelligibility, then, high-*R* words are "easy" and low-*R* words are "hard". It is this sort of listener difficulty in word recognition that will be further explored in the current study.

(1) High vs. Low Relative Frequency: Target words are represented by dark bars, neighbors by light bars; frequency is represented by bar height.



That this notion of lexical confusability, measured by relative frequency, represents a real perceptual or intelligibility factor is verified by auditory lexical decision experiments showing faster response times to "easy" (high-*R*) words than to "hard" (low-*R*) words, indicating that "easy" words are more easily processed (Luce 1986; see also the current study). Such a measure also represents a real factor in production, as evidenced by the finding that "easy" words are produced with consistently greater vowel reduction than "hard" words (Wright 1997). In other words, the less easily processed words (the "hard" ones) are produced with more careful, acoustically distinct vowels, increasing the likelihood of intelligibility in precisely those cases where processing is more difficult (as evidenced by slower reaction times) and where a listener might otherwise be likely to misperceive.

3. Coarticulatory Modifications

The current study looks at another way speakers might accommodate the needs of listeners, investigating whether coarticulation is mediated by potential lexical confusability.

There are two competing hypotheses regarding the effect of lexical confusability on coarticulation. On the one hand, coarticulation introduces variability into the speech signal. Coarticulation is often viewed as the primary means by which speakers reduce their articulatory effort. But as coarticulation reduces speakers' effort, if it does, it may also reduce acoustic distinctiveness and hurt the ability of listeners to accurately or easily perceive the speech signal. From the point of view of the listener, then, coarticulation is just variability or noise that must be factored out of the signal in order to understand a speaker's message. Thus, in cases where speakers try to accommodate the difficulties of their listeners, it is predicted that they would increase their effort and decrease coarticulation, particularly in "hard" words.

On the other hand, though coarticulation may lead to a reduction of articulatory effort, it does not actually lead to a reduction of the amount of acoustic information contained in the speech signal. Coarticulation involves the overlapping of segments, spreading acoustic properties of one segment to another segment or segments. But these acoustic properties from different segments are not merged into a single, ambiguous acoustic stream. Rather, they are all present in the signal and available to listeners to perceive (Mattingly 1981). If the cues present in coarticulation can be used by listeners, then acoustic enhancement of certain segments by coarticulation could be beneficial to listeners in perceiving words. These considerations lead to the prediction that a speaker who is trying to be maximally intelligible would increase the degree of coarticulation in his speech, especially in "hard" words.

The current study investigates the effect of the listener-oriented constraint of lexical confusability on the production of coarticulation to determine whether coarticulatory modifications are used to increase speaker intelligibility in listener-directed speech. First, a production experiment seeks to resolve which of the two possible coarticulatory modifications hypothesized above, an increase or a decrease in coarticulation (if either), occurs in listener-directed speech. Then, a perception experiment tries to determine whether the coarticulatory modifications made in lexically confusable words is actually beneficial to listeners.

4. Production Experiment

Speakers were recorded producing words in two categories, high-*R* ("easy") and low-*R* ("hard"), which exemplify two types of coarticulation, carryover vowel-to-vowel coarticulation and anticipatory nasal coarticulation. The degree of coarticulation was then measured and compared across lexical confusability categories. The prediction was that coarticulation would be either reduced or exaggerated in "hard," or low-*R*, words relative to high-*R* words, reflecting speakers' knowledge about the effect of coarticulation on intelligibility.

4.1. Methods

4.1.1. Speech Materials

Tokens in the study included 40 monosyllabic words with codas containing nasals and 40 disyllabic words of the form CVC/i/. Within each set of tokens, half of the words were high-*R* and half were low-*R*. Lexical frequencies and the set of neighbors were determined based on the CELEX corpus (with adjustments for American English pronunciation) (Baayen, et al. 1995). For the purpose of the study, neighbors were considered to be all words differing from the target word by the addition, deletion, or substitution of a single phoneme (Greenberg & Jenkins 1964; Luce 1986). Relative frequency (*R)* was calculated by dividing the $\log_{10}$ frequency of the token word by the sum of the $\log_{10}$ frequencies of the word and all of its neighbors. Words were selected for inclusion in the study on the basis of having *R* values at the upper or lower end of the range of *R* values for words of the same type (i.e., CVNs or CVC/i/s). Additionally, mean log frequency and segmental context were balanced across the two relative frequency conditions for each type of coarticulation, and all words were of approximately equal, high subjective familiarity (Nusbaum, et al. 1984). For the purpose of comparison with the coarticulated vowel in the CVC/i/ words, a set of seven words exemplifying a canonical, non-coarticulated /i/ were also included in the list of tokens to be recorded. Four of these words were of the form C/i/C with no context for V-to-V coarticulation, and three were of the form C/i/C/i/ where the /i/s could only be coarticulated with themselves, yielding a form free from coarticulatory influence from other vowels. Two different randomized lists of the test words were created. The canonical /i/ words were included in a set of 12 practice words. (See Brown (2001) for a complete word list.)

4.1.2. Procedure

To encourage a listener-directed speech style, subjects participated in recording sessions in pairs, taking turns acting as speaker and listener. The speaker saw each word, one at a time, on a computer screen (not visible to the listener) and dictated a list to the listener, saying "The first word is X. The word after X is Y", etc. The speaker controlled the rate of presentation by pressing a button to advance to the next word. The procedure began with 12 practice tokens and continued until all 80 test tokens were uttered by the speaker. The roles of the two subjects were then reversed so that the subject who had been the speaker became the listener, and vice versa. The second speaker presented the words to the listener in the other random order. The procedure yielded two repetitions of each word for each speaker: one in utterance-final position, and the other in the "word after X" context. The data of seven native speakers of American English are analyzed here.

4.1.3. Measurements

The degree of nasal coarticulation in the CVN words was determined by measuring the nasality of the nasalized vowel from the acoustic signal. Nasalized vowels show the presence of an extra spectral peak at low frequencies, generally below the first formant, as well as a reduction in the amplitude of the first formant spectral peak. A measure of the relative amplitudes of the nasal peak and the first formant, A1-P0, where A1 is the amplitude of the first formant (as estimated by the amplitude of the peak harmonic closest to the expected F1) and P0 is the amplitude of the low frequency nasal peak (as estimated by the amplitude of either the first or second harmonic, depending on the speaker), has been found to be a reliable quantification of vowel nasality (Chen 1997). A1 and P0 were measured at the vowel midpoint and at the end of the vowel.

The degree of vowel-to-vowel coarticulation in the CVC/i/ words was determined by comparing the F1 and F2 of the second vowel (/i/) with the mean F1 and F2 values of the canonical, non-coarticulated /i/s for the same speaker. Since /i/ is higher and more front than any of the vowels which might be coarticulated with it, coarticulatory influence from any vowel would be realized as a raising of F1 and/or a lowering of F2. The first two formants of the canonical /i/ were determined by averaging the F1 and F2 of 18 /i/s in the practice words with no context for V-to-V coarticulation. Because V-to-V coarticulation is most visible in the region of the second vowel (V_2) closest to the first vowel (V_1), measurements for all vowels were made at V_2 onset and 20 ms into the vowel. All frequencies in Hertz were converted to the auditorily-scaled Bark scale, since the current study is concerned with the auditory salience of any acoustic differences that are found. Additionally, the transformation allows for the direct comparison of F1 and F2 differences. Differences between canonical /i/ and coarticulated /i/ were then calculated for both F1 and F2 at each measurement point.

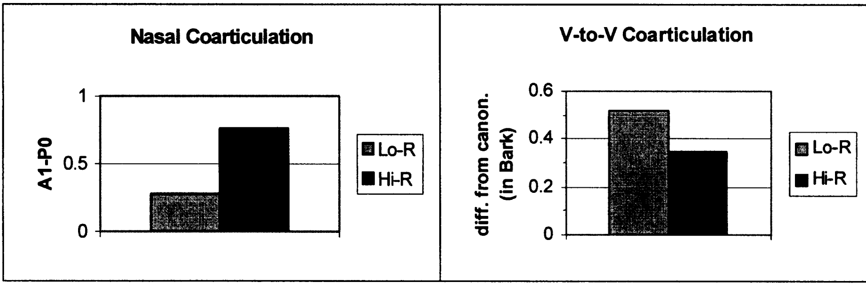
4.2. Results

Mixed repeated measures ANOVAs with nested within-subject factors of *R* (High vs. Low) and *Repetition* (First vs. Second) and a between-subjects factor of *RoleOrder* (A or B, indicating whether the subject was the first or second speaker in his session, with list randomization A assigned to the first speaker and B always assigned to the second) were performed separately on the A1-P0 values of the CVNs and the F1 and F2 differences of the CVC/i/s (averaged within each test category, for each measurement position).

For the nasal coarticulation items, a main effect of *R* was found at the vowel midpoint [$F(1,5)=11.77$, $p=.019$], reflecting that low-*R*, or "hard," words, have more nasalized vowels (i.e., lower A1-P0 values) than high-*R*, or "easy," words at the midpoint. This effect of *R* on nasality is summarized in the first graph in (2) below. No other main effects or interactions reached significance at any position, though low-*R* words also show smaller A1-P0 values than high-*R* words for five of the seven subjects at the end of the vowel. It was further observed that the A1-P0 values are smaller at the end of the vowel than at its midpoint, indicating that the vowels were more nasal closer to the nasal consonant.

For the V-to-V coarticulation items, there was again a significant main effect of *R* on F1 at V_2 onset [$F(1,5)=21.41$, $p=.006$], reflecting that the final vowels in low-*R* ("hard") words differ more from canonical /i/, presumably due to influence from V_1, than those in high-*R* ("easy") words. This effect of *R* is shown in the second graph in (2). No other main effects or interactions reached significance at V_2 onset or at 20 ms after onset, though F1 means at 20 ms show the same *R* pattern as that found at vowel onset. The F2 differences across high- and low-*R* words do not appear to vary systematically, however.

(2) Left: A1-P0 for Low-*R* vs. High-*R* CVN Words, at vowel midpoint. Right: F1 Difference between Canonical /i/ and Experimental /i/ for Low-*R* vs. High-*R* CVC/i/ Words, at vowel onset.



4.3. Discussion

The greater degree of nasality in low-*R* words relative to high-*R* words among CVNs indicates greater anticipatory nasal coarticulation in "hard", or more lexically confusable, words. Likewise, the greater deviation from canonical vowel values for F1 (at onset) indicates greater carryover coarticulation (at least on the height dimension) for "hard" CVC/i/ words. Thus, we see that "hard" words are produced with *more* coarticulation than "easy" words.

If we take the view that coarticulation is a sort of reduction (at very least, a strategy by which speakers reduce their overall effort) (e.g., Lindblom 1990), these results are quite surprising. The object of listener-directed speech is supposed to be to make speech more clear, i.e., less reduced, when there are greater demands on the listener. Wright's (1997) finding that "hard" words are produced with more dispersed vowels than "easy" words seems to underscore this principle of enhancing acoustic distinctiveness and minimizing reduction with respect to the demands of lexical confusability.

However, because the results of the current study indicate that speakers increase coarticulation in precisely those words that they otherwise try to make clearer, they suggest a very different view of coarticulation in which coarticulation is not in conflict with listeners' need to receive clear, distinct acoustic information. While coarticulation may in a sense diminish the acoustic distinctiveness of one segment (in that it makes the segment on which it is realized less canonical), it provides more extensive cues for another. It appears, then, that speakers produce low-*R* words with more coarticulation so that listeners can use the extra information contained in the coarticulation to help them perceive these "hard" words more easily.

If this story is correct, and coarticulation is increased in "hard" words in order to make these words maximally intelligible, gaps in the effect at certain measurement points or on certain dimensions are not expected. No effect was found, though, for nasality at the vowel endpoint in the case of nasal coarticulation or for F2 in the V-to-V case. Before we continue, then, we will briefly consider these cases. It is important to note that in these two instances, there is not a lack of coarticulation; in fact, there is as much coarticulation on the F2 dimension as there is for F1 and more nasal coarticulation at the endpoint than is found at other positions. What is lacking is simply an influence of *R*.

Looking first at A1-P0 at the endpoint in CVNs, it may be noted that, across all tokens, A1-P0 is significantly smaller at this point than at the midpoint [$t(28)=2.052$, $p<.0001$], indicating that the end of the vowel is more nasalized.

The lack of an effect of *R* can be straightforwardly interpreted as a ceiling effect. If the velum consistently reaches a fully low position before the closure for the nasal stop (toward the end of the vowel), it is not possible (or necessary) to differentiate the nasality of vowels in low-*R* words and high-*R* words by lowering the velum even further in the low-*R* words, since the vowels in both would already be maximally nasalized, produced with a maximally lowered velum.

With respect to F2 in the CVC/i/s, it may be noted that at both measurement positions, F2 differs as much from canonical in low-*R* ("hard") words as F1 does, suggesting the same amount of coarticulation for F1 and F2 in low-*R* words. However, there is also an F2 deviation from canonical in the high-*R* ("easy") words that is equal to that in the low-*R* words. It is possible, of course, that there is simply as much F2 coarticulation in "easy" words as in "hard" ones, even though we would not expect speakers to make any effort to adjust coarticulation in "easy" words. An F2 deviation, however, might also be attributable to reduction, an especially likely possibility for "easy" words, which undergo greater vowel reduction than "hard" words (Wright 1997). If F2 in "hard" words shows *R*-driven coarticulation, and F2 is lowered in the "easy" words due to vowel reduction, these two effects on F2 may balance each other out in the statistical analysis, hiding the effect of *R*. (Because there are more height (F1) contrasts than front-back (F2) contrasts to maintain in English, F1 is less likely to reduce than F2.) It would be helpful to examine the relative effects of reduction on F1 and F2 in "easy" words without a potential coarticulatory confound, though data on the contribution of the formants individually are not currently available.

5. Perception Experiment

We have supposed thus far that if speakers pronounce certain sounds differently depending on the relative frequency (*R*) of the word in which they occur, they must be doing so for the benefit of listeners, to make the words which are potentially the most difficult to perceive as easy as possible for the listener to understand. The purpose of the second experiment, then, is to investigate whether differences in pronunciation due to adjustments of the degree of coarticulation actually affect a listener's ability to perceive words. In the current experiment, listeners heard real words and nonsense words of two phonetic types: those which displayed normal coarticulatory effects and those in which coarticulatory effects were removed. Lexical decision reaction times, indicating a listener's facility in perceiving a word, from these two coarticulation conditions were compared. Based on the findings for low-*R* words in the production study, it was predicted that listeners would respond more quickly to tokens *with* coarticulation than to those *without*.

5.1. Methods

5.1.1. Stimuli

The set of stimuli included items that fall into three groups: test words, filler words, and non-words. The tokens of primary interest were the real test words, which were the same 40 CVNs and 40 CVC/i/s used in the production experiment. The set of filler words included 20 monosyllabic and 20 disyllabic forms, none of the form CVN or CVC/i/. The non-words included those corresponding to the test words, which were formed by changing the onset of the initial syllable, and those corresponding to the fillers, which were formed by changing any syllable onset or coda.

Each of the stimuli presented to subjects was produced by splicing together parts of two separate recordings. In some cases, the two recordings were different productions of the same item (referred to as "same-spliced"); in other cases, they were recordings of different items (referred to as "cross-spliced"). For each of the real test words, both a same-spliced version and a cross-spliced version were made. For the remaining items, only one version, either same-spliced or cross-spliced was made. The original components of the stimuli were produced by a female native speaker of American English. The speaker read a list containing the test words, fillers, non-words, and tokens resembling these items except that the segment inducing the relevant coarticulation was replaced, eliminating the context for coarticulation. For example, for the CVC/i/ item *chubby*, both *chubby* and *cheeby* were recorded. The tokens were spliced as shown in the figure in (3).

(3) Illustration of Splicing Scheme for V-to-V and Nasal Coarticulation

		Same-spliced		Cross-Spliced	
V-to-V	Originals	**tʃʌb$_{ʌ}$i**	**tʃʌb$_{ʌ}$i**	**tʃʌb$_{ʌ}$i**	**tʃib$_{i}$i**
	Spliced stimulus	**tʃʌb$_{ʌ}$i**		**tʃʌb$_{i}$i**	
Nasal	Originals	**bʌ̃n**	**bʌ̃n**	**bʌ̃n**	**bʌd**
	Spliced stimulus	**bʌ̃n**		**bʌn**	

Because the reaction times to the pairs of real test words were to be compared (same-spliced versus cross-spliced), special care was taken to ensure that members of each pair were as similar as possible. The same sound file was used as the skeleton (the word without the critical vowel) for both versions, and the durations of the critical vowels were matched as closely as possible across the versions, differing by no more than 4 ms. (Further details on stimulus preparation are available in Brown (2001).)

All subjects heard the same 240 words and non-words; however, subjects were randomly divided into two groups with respect to which versions (same-spliced or cross-spliced) they heard of each word. Listeners in each group heard half of the real test words in each version, but the items that were heard in one version by one group were heard in their other version by the other group. Thus, each test token (i.e., a word in a particular splicing condition) was heard by half the subjects. The stimuli were heard in a random order.

5.1.2. Procedure

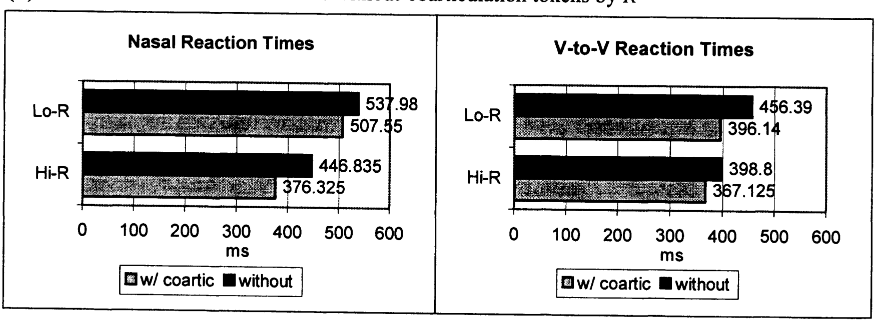
Subjects were instructed that they would hear a number of simple English words and some nonsense words and that their task was to decide quickly and accurately which were real words and which were non-words and to press a response button on a button box to indicate their choice. New trials were initiated 1 second after each button press. After 12 practice trials, responses ("word" or "non-word") and response latency times were recorded for 240 trials. Nineteen native speakers of American English, none of whom had participated in the first experiment, participated in this experiment.

5.2. Results

Overall, 91% of the 1520 trials involving the relevant real test words were correctly identified as words. Accuracy did not differ significantly between same- and cross-spliced tokens. In the analysis of reaction times, only data from real words that were correctly identified were considered. Furthermore, response times that did not fall within two standard deviations of the mean for each subject were also removed from the analysis.

Reaction times (averaged within each condition) for each subject were submitted to mixed repeated measures ANOVAs (one for each type of coarticulation, V-to-V and nasal) with nested within-subjects factors of Coarticulation (with vs. without, represented by the same- vs. cross-spliced tokens) and *R* (high vs. low) and a between-subjects factor of List (A or B, indicating which set of words subjects heard in which condition). For both types of coarticulation, there is a significant main effect of *R* [Nasal: $F(1,17)=61.57$, $p<.0001$], [V-to-V: $F(1,17)=16.7$, $p=.0008$], reflecting that low-*R* words were, as expected, harder (i.e., they took longer) for subjects to verify. The hypothesized main effect of Coarticulation was also found to be significant for both coarticulation types [Nasal: $F(1,17)=28.54$, $p<.0001$], [V-to-V: $F(1,17)=5.77$, $p=.028$], reflecting the fact that words *with* coarticulation (i.e., same-spliced words) were more quickly verified than words with no coarticulation (i.e., cross-spliced words). The reaction time data are summarized in (4).

(4) Reaction Times to with- vs. without-coarticulation tokens by *R*



5.3. Discussion

The results from the lexical decision experiment showing a significant effect of Coarticulation support the hypothesis that the presence of coarticulation facilitates lexical perception. Interestingly, these results obtain both when coarticulation provides cues for an upcoming segment (as in the nasal case) *and* when it reinforces cues following a segment (as in the carryover V-to-V case).

Earlier studies have shown that listeners are able to attend to coarticulatory information when they are asked to make explicit, fine, phonetic discriminations (e.g., Beddor, et al. 2001), and that they are sensitive to the inappropriateness of certain bits of coarticulatory information (e.g., Fowler 1984). The current study extends these findings, allowing for broader generalizations regarding listeners' abilities to use information contained in coarticulation. First, it examined a relatively more linguistically-natural task. Listeners' reactions to appropriately coarticulated strings were compared to coarticulatorily-neutral strings, rather than to instances of inappropriate coarticulation. Because inappropriate coarticulation (e.g., vowel nasality before an oral consonant) is clearly unnatural, it might be expected to be harder to respond to on the basis of a general listener dispreference for unnaturalness, whether or not the presence of appropriate coarticulatory cues in the other condition contributed to improved reaction times. In the current study, however, the inappropriate coarticulation condition was replaced with a neutral condition, in which coarticulation was simply removed. Because coarticulation naturally varies in degree, tokens that lack coarticulation should not sound completely unnatural; rather they should sound like tokens produced with the smallest possible degree of coarticulation. Furthermore, the current study looked at listeners' reactions to coarticulation as they were processing real words (as opposed to nonsense phonetic strings). The current findings, then, suggest that listeners are not only able to extract phonetic information from coarticulation, but that they actually do make use of coarticulatory information (at least under the circumstance of timed lexical decision) when they perceive real words.

6. Summary and General Discussion

The purpose of the current study was to investigate the effect of potential lexical confusability on the production of coarticulation. We asked whether speakers would try to accommodate listeners' difficulties in spoken word recognition caused by lexical confusability by adjusting the amount of coarticulation in particularly confusable words. And we looked at the effect of the presence of coarticulation on lexical identification to see whether the accommodations made by speakers would actually be beneficial to listeners. It was found that more lexically confusable words, those with low frequencies relative to their lexical neighbors, were produced with a greater degree of coarticulation than less confusable words for both types of coarticulation examined (carryover vowel-to-vowel and anticipatory nasal coarticulation). It was also shown that listeners were able to more quickly identify words containing coarticulation than words in which the coarticulation had been eliminated.

Taken together, the results from these two experiments provide a coherent story within the framework of listener-directed speech accommodations. Listeners are able to recognize words with more coarticulation more quickly and easily than those without, so speakers adjust their pronunciation of harder, more confusable words to exhibit a greater degree of coarticulation in order to facilitate lexical recognition. To explain this phenomenon, coarticulation can be thought of as a process that spreads the properties of a segment, causing them to overlap in time with cues from neighboring segments. (For a discussion of this sort of spreading process in phonology, see Kaun (1995) in which certain instances of phonological assimilation are characterized as spreading to enhance features.) Listeners are able to compensate for coarticulation, correctly attributing the acoustic consequences of coarticulation to their source (e.g., Beddor & Krakow 1999; Mann & Repp 1980). And they are also able to use coarticulatory information to identify or predict other portions of the signal (e.g., Fowler 1984; Beddor & Krakow 1999). Thus, increased coarticulation in "hard" words (at least to a certain degree) can provide extra cues for the coarticulatory source segment without interfering with the cues for the segment on which they occur. This is the effect that emerges from the present study.

The current study, then, suggests that speakers have some knowledge of lexical confusability or relative frequency that leads them to identify certain words as potentially difficult (with respect to lexical access) and that triggers them to make various communicatively-motivated accommodations. It demonstrates that one such accommodation involves an increase in the degree of coarticulation in "hard" words. And it suggests that coarticulation is more than just a type of reduction that allows speakers greater ease of articulation. Some coarticulation (the amount that we see in high-*R* words) may in fact facilitate articulation (and may even be necessary for articulation). But beyond this, coarticulation (for example, in the amount that we see in low-*R* words) serves an explicitly listener-oriented, communicative purpose, providing additional cues that facilitate lexical perception.

References

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995) The CELEX Lexical Database (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [distributor].

Beddor, P. S. & Krakow, R. A. (1999) Perception of coarticulatory nasalization by speakers of English and Thai: evidence for partial compensation. *JASA*, 106, pp. 2868-2887.

Beddor, P. S., Krakow, R. A., & Lindemann, S. (2001) In E. V. Hume & K. Johnson (Eds.), *The Role of Speech Perception Phenomena in Phonology.* Academic Press.

Brown, R. A. (2001) *Effects of Lexical Confusability on the Production of Coarticulation.* UCLA master's thesis.

Chen, M. Y. (1997) Acoustic correlates of English and French nasalized vowels. *JASA*, 102, pp. 2360-2370.

Fowler, C. A. (1984) Segmentation of coarticulated speech in perception. *Perception & Psychophysics*, 36, 359-368.

Fowler, C. A. & Housum, J. (1987) Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. *JML*, 26, 489-504.

Glanzer, M. & Eisenreich, S. L. (1979) Structure and search of the internal lexicon. *J. Verbal Learning and Verbal Behavior*, 18, 381-398.

Gordon, B. (1983) Lexical access and lexical decision: mechanisms of frequency sensitivity. *J. Verbal Learning and Verbal Behavior*, 18, 24-44.

Greenberg, J. H. & Jenkins, J. J. (1964) Studies in the psychological correlates of the sound system of American English. *Word*, 20, 157-177.

Hanley, T. D. & Steer, M. D. (1949) Effect of level of distracting noise upon speaking rate, duration and intensity. *J. Speech and Hearing Disorders*, 14, 363-368.

Jakobson, R., Fant, C. G. M., & Halle, M. (1952) Preliminaries to speech analysis: the distinctive features and their correlates. *Technical Report, no. 13.* Cambridge, MA: Acoustics Laboratory, MIT.

Kaun, A. R. (1995) *The Typology of Rounding Harmony: An Optimality Theoretic Approach.* UCLA dissertation.

Lieberman, P. (1963) Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172-187.

Lindblom, B. (1990) Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modeling*, pp. 403-439. Netherlands: Kluwer Academic Publishers.

Lombard, E. (1911) Le signe de l'élévation de la voix. *Annales des maladies de l'oreille, du larynx, du nez et du pharynx*, 37, 101-119.

Luce, P. A. (1986) Neighborhoods of words in the mental lexicon. *Research on Speech Perception Progress Report, No. 6.* Speech Research Lab., Psych. Dept., Indiana University.

Luce, P. A. & Pisoni, D. B. (1998) Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*, 19, 1-36.
Mann, V. & Repp, B. (1980) Influence of vocalic context on perception of the [š]-[s] distinction. *Perception & Psychophysics*, 28, 213-228.
Mattingly, I. G. (1981) Phonetic representation and speech synthesis by rule. In T. Meyers, J. Laver, & J. Anderson (Eds.), *The Cognitive Representation of Speech,* pp. 415-420. Amsterdam: North-Holland Publishing Company.
Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984) Sizing up the Hoosier mental lexicon: measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, No. 10.* Speech Research Lab., Psych. Dept., Indiana University.
Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985) Speaking clearly for the hard of hearing I: intelligibility differences between clear and conversational speech. *J. Speech and Hearing Research*, 28, 96-103.
Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986) Speaking clearly for the hard of hearing II: acoustic characteristics of clear and conversational speech. *J. Speech and Hearing Research*, 29, 434-446.
Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985) Speech perception, word recognition, and the structure of the lexicon. *Speech Communication*, 4, 75-95.
Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988) Effects of noise on speech production: acoustic and perceptual analyses. *JASA*, 84, 917-928.
Wright, R. (1997) Lexical competition and reduction in speech: a preliminary report. *Research on Spoken Language Processing Progress Report, No. 21.* Speech Research Lab., Psych. Dept., Indiana University.

Rebecca Scarborough
3125 Campbell Hall
UCLA Linguistics Department
Los Angeles, CA 90095

rab@humnet.ucla.edu

On the Syntactic Status of Certain Fronted Adverbials in English*

BENJAMIN SHAER
Zentrum für Allgemeine Sprachwissenschaft, Berlin

0. Introduction

In recent years, there has been a great deal of interest in the syntactic, semantic, and discourse properties of the sentence's left periphery. Despite this interest, many questions about this domain remain open. Among them is a question to be addressed in this paper, specifically as it pertains to English. This is whether fronted adjuncts—those displaced to the left periphery from their canonical positions in the sentence—are better analysed as base-generated there or extracted from some lower base position. As it happens, evidence has been accumulating that *wh-* and other left-peripheral (LP) adjuncts across languages are indeed extracted. Hukari and Levine (1995), for example, assemble an array of data supporting the view that *wh*-adjuncts are extracted rather than generated 'in situ'. Most relevant here are their data showing crossover effects among English *wh*-adjuncts, as illustrated below, where such effects can be understood to involve the illicit extraction of the adjunct over a coindexed element that either is or is not embedded in a larger constituent, corresponding to weak and strong crossover, respectively:

(1) a. * [In whose$_i$ bedroom]$_j$ did Mary claim his$_i$ mother asked Sarah to have lunch e$_j$? (= weak crossover)
    b. * [In whose$_i$ bedroom]$_j$ did Mary claim he$_i$ asked Sarah to have lunch e$_j$? (= strong crossover)

Sentences like these give us good reason to believe that the preposing of *wh*-adjuncts in English is subject to syntactic constraints, and thus that these cases

* This paper represents a development of work reported in Shaer (2003a: §4.3) and shares some content with papers presented at the meetings of the West Coast Conference on Formal Linguistics in San Diego, the Linguistics Association of Great Britain in Sheffield, and the North American Syntax Conference in Montreal. I wish to thank the audiences of BLS 29 and these other conferences and Werner Frey, Nikolas Gisborne, and Claudia Maienborn for helpful discussion.

should be treated on a par with argument extraction (see Hukari and Levine 1995: 222).

In frameworks that take the extraction of an element to involve movement from a base to a target position, these and other preposing data are naturally treated in terms of the claim that movement is not a free option in the grammar, but must instead 'be triggered by the satisfaction of certain quasi-morphological requirements of heads', which are related to particular features of these heads (Rizzi 1997: 282). Since such 'forced' movement requires 'the presence of a head entering into a Spec-head configuration with the preposed phrase', it accordingly requires an 'articulated structure to host the different kinds of phrases moved to the left periphery' (Rizzi 1997: 282)—hence the 'articulated C' approach to this domain advocated by Rizzi and others. On such an approach, topicalized and focalized constituents as well as *wh*-phrases move to the specifiers of a cascade of LP functional projections, the heads of these projections encoding information related to the discourse function of these moved constituents and the sentence's illocutionary force or clause type. This approach can thus claim among its advantages its ability not only to extend a single movement mechanism to a range of A′-movement structures, but also to capture observed ordering patterns among preposed arguments and adjuncts and to articulate the widely held view that certain left-peripheral positions are closely tied to certain discourse functions and to the encoding of a sentence's clause type.

Yet doubts about this approach have been raised by many researchers. These doubts have centred on the difficulties that face functional structure accounts of ordering effects and the problematic nature of feature-driven topic and focus movement and of the purported connection between syntactic position and discourse function.

A rather different sort of argument against such uniform treatments of adjunct preposing comes from consideration of the work of Espinal (1991) and others, who have claimed that English LP discourse adverbials such as *frankly* are base-generated as 'orphans', elements that are 'not syntactically integrated [into the sentence] at any level of representation' (Haegeman 1991). It is this sort of argument that I shall be pursuing in what follows. What I shall show is that English LP adverbials exhibiting the properties of orphans constitute a class considerably larger than that recognized in the literature, and one that includes LP manner and instrumental adverbials as well as discourse adverbials. Moreover, evidence that these adverbials are orphans rather than moved constituents suggests that languages may structure their left periphery in significantly different ways, creating a picture of cross-linguistic variation that goes beyond the outlines sketched by the 'articulated C' approach. Taken together with other doubts, such considerations point away from this approach and toward a 'modular' approach to the left periphery, in which the discourse properties of constituents and the force of sentences are not determined directly on the basis of syntactic properties, but rather through the interaction of the gross structural properties of sentences with lexical, phonological, and other linguistic properties of expressions, giving rise to the complex patterns that we observe.

The rest of this paper is organized as follows. First, I shall briefly review the 'articulated C' approach (section 1). Next, I shall describe some LP adverbials that, despite first appearances, are not plausibly analysed as moved (section 2), and then show that these might not even be in a c-command relation to the rest of the sentence—the heart of the 'orphan' claim (section 3). Then I shall address some open questions about orphans and their compatibility with an 'articulated C' approach (section 4). Finally, I shall offer some concluding remarks (section 5).

1. The 'articulated C' analysis of the left periphery

A common claim in the literature is that the sentence's left periphery is organized into a syntactic 'topic' domain into which elements are moved, certain key discourse properties of these elements thus being encoded in their syntactic positions. As noted above, a recent family of analyses has cashed out this claim in terms of a 'C domain' composed of various functional projections into which topicalized and focalized elements move to satisfy specifier-head requirements. According to the best-known version of this analysis, that of Rizzi (1997), the C domain has the following structure. The highest projection in the domain is Force Phrase, which encodes the illocutionary force or clause type of the sentence, the head of which may be occupied by subordinating conjunctions. Below this projection are iterating Topic Phrases that appear above and below Focus Phrase, the specifiers of these projections hosting topicalized and focalized constituents, respectively. Below the lower Topic projection is Finite Phrase, which specifies the finiteness of the selected IP (Haegeman 2002: 16). This structure is schematized in 2, and the kind of sentences taken to motivate it exemplified in 3:

(2) Force > Top* > Foc > Top* > Fin (Rizzi 1997: 288, (8))

(3) Credo che a Gianni, QUESTO, domani, gli dovremmo dire
C Top Foc Top IP
'I believe that to Gianni, THIS, tomorrow we should say.'
(Rizzi 1997: 295, (37a))

More recent work on the C domain has proposed various modifications to this structure, involving further functional distinctions and loci of variation both between languages and between clause types within a language. For example, Rizzi (2002) and Haegeman (2002) add the lower projection Mod(ifier) to host adjuncts, in order to account for the inability of fronted adjuncts to occur to the left of elements in Focus; and Haegeman (2002) adds a Sub(ordinator) position, to distinguish between 'the head that encodes "force" (clause type) and the head that serves simply to subordinate the clause'. In addition, Haegeman (2002) argues for two kinds of variation in 'articulated C' structure: the absence of a lower Topic position in English and the existence of 'truncated' structures for certain kinds of adverbial clauses, from which Force, Top, and Focus are absent. These revised C structures are given in 4:

(4) a. Sub > Force > Top* > Focus > Mod* > Fin > IP
b. Sub > Mod > Fin > IP (Haegeman 2002: §7.1, (69))

Now, while this view of the 'C domain' has gained many adherents, it remains a controversial one, about which many authors have already expressed scepticism. Newmeyer (2002), for example, has argued that certain ordering patterns taken to follow from an articulated C domain have an independent explanation. Among these is the pattern given in 5, which indicates that a relative pronoun may occur with a topicalized argument only when the former is higher than the latter:

(5) a. He's a man to whom liberty we could never grant.
(Baltin 1982: 17, (69))
b. * He's a man, liberty, to whom we could never grant.

This pattern is readily captured on an articulated C analysis in terms of the fixed positions of the two constituents in question, which occupy the specifiers of ForceP and TopP, respectively. Yet according to Newmeyer, this pattern can be accounted for in terms of an independently required Nested Dependency Constraint, whereby 'multiple filler-gap dependencies may be disjoint or nested, but not intersecting' (Newmeyer 2002). If Newmeyer is correct, then the account of such patterns provided by an articulated C analysis may be dispensed with on the grounds of parsimony. Moreover, the appeal to a specifier of ForceP position for relative pronouns, though necessary to account for the fact that these elements occupy a position higher than the topicalized NP, is nevertheless an implausible one given its implications for the encoding of illocutionary force. This is simply because relative pronouns have no obvious relation to a sentence's illocutionary force, and are thus not naturally treated as hosted by this projection.

Newmeyer points out similar difficulties for the claim that focus is associated with a particular syntactic position. These include the fact that focus may be associated with formally discontinuous elements or even bound morphemes, as the sentences in 6 show:

(6) a. A: Did Mary wash the car? B: No, TOM washed the WINDOWS.
b. John is more concerned with AFfirmation than with CONfirmation.

In addition to these doubts about the relation between syntactic position and discourse function claimed by articulated C analyses are those associated with their claims about the syntactic encoding of illocutionary force. The brief remarks that we find on this matter—for instance, Rizzi's (1997: 283) claim that complementizers, as the heads of ForceP, 'express the fact that a sentence is a question, a declarative, an exclamative, a relative, a comparative, an adverbial of a certain kind, etc.'—suggest a conception of illocutionary force that bears only a passing resemble to this notion as it is generally understood. In particular, such a description overlooks the fact that illocutionary force irreducibly involves speaker intention. As Zanuttini and Portner (to appear: §1) note, a sentence has, for example, 'the illocutionary force of ordering if and only if the speaker intends to impose an obligation by getting the hearer to recognize this intention' regardless of the particular form of the sentence. Thus, an order may just as easily have the grammatical form of a question as that of an imperative, as indicated in 7:

(7) a. Could you come in at 9:00?
 b. Come in at 9:00!

On this basis, Zanuttini and Portner (to appear: §1) argue that illocutionary force is 'not the right sense of force for the characterization of clause types', and that what is relevant is 'sentential force', 'the force conventionally associated with the sentence's form'. Yet, as they show, even this more restricted notion of force cannot plausibly be seen as 'syntactically realized as a single element or feature', or as 'consistently associated with a single projection'.

As the above remarks suggest, there are many reasons to doubt that complex ordering effects, discourse function, or illocutionary force can be insightfully analysed in terms of a cascade of functional projections in the C domain. A further reason to doubt the explanatory utility of this approach, which we turn to in the next section, is related to its claim that phrases assume the roles of 'topic' and 'focus' by moving to positions in the tree designated for these discourse functions. What we shall see there is evidence that certain fronted adverbials in English are not moved.

2. Some 'lower' LP adverbials that are not moved

Among the fronted adverbials that, I shall argue, are not plausibly treated as moved are those in 8:

(8) a. With perfect technique, John executed the triple somersault.
 b. As passionately as he could, Kim kissed Sandy.
 c. With nothing but a crowbar and a ballpeen hammer, I very much doubt that Terry will be able to repair the Vax in our office.
 (Hukari and Levine 1995: 201, (16a))

What we find here are adverbials that receive manner and instrumental readings—that is, ones generally associated with VP-internal base positions. Yet these adverbials differ in various respects from those that are uncontroversially moved, such as the ones in 9:

(9) a. Only very quietly did John get drunk.
 b. How quietly did John get drunk?

Among these differences is that the fronted adverbials in 8 but not 9 are set off intonationally from the rest of the sentence;[1] and those in 9 but not 8 are associated with subject-auxiliary inversion. There also appear to be key differences in the behaviour of these two sets of adverbials. One of particular interest, for reasons that will emerge presently, is that the adverbials in 8 do not give rise to Principle C violations when they contain R-expressions coindexed with a subject NP, whereas those in 9 do, as indicated in 10 and 11, respectively:

(10) a. As fast as John's$_i$ little legs could carry him, he$_i$ ran to warn the others.
b. More loudly than Mary$_i$ had hoped, she$_i$ tiptoed along the dark corridor.
c. With nothing but a crowbar and Terry's$_i$ ballpeen hammer, I very much doubt that he$_i$ will be able to repair the Vax in our office.

(11) a. * So much faster than John's$_i$ personal best was he$_i$ running that the dog got scared.
b. * Only with Ben's$_i$ bare hands could he$_i$ strangle the chicken.

This contrast is an intriguing one, since it is not predicted by any account that takes the adverbials in 10 to be A′-moved to positions high in the tree and reconstructed to their original positions. This suggests that these adverbials have not moved, whereas those in 11 have.

Intriguing though this contrast may be, enough difficulties have been raised regarding the status of Principle C in the grammar (e.g. Pollard and Sag 1994: 247–48; Shiobara 2003) to encourage us to seek out further evidence that adverbials like those in 10 are not moved.

As it happens, other evidence is indeed available. This includes the failure of the adverbials in question to obey standard movement constraints, as indicated by the contrast between the (a) and (b) sentences below:

(11) a. With his 'spider powers' alone, I very much doubt Peter's boast that he can wipe out terrorism.
b. * With which of his$_i$ special powers alone did Mary question Peter's$_i$ boast that he could wipe out terrorism?

(12) a. Even using 'spider powers', you wonder who could wipe out terrorism.
b. * By what means exactly do you wonder who could wipe out terrorism?

[1] Although, as Taglicht (2001: 4, 6) points out, this matter turns out to be more complicated, since these adverbials, like their parenthetical counterparts, are typically but not uniformly associated with prosodic boundaries.

This evidence also includes the inability of these adverbials to occur above sentence adverbs like *likely* and *probably,* as in 13a. This behaviour again contrasts with that of adverbials associated with subject-auxiliary inversion, as 13b indicates, suggesting that the former adverbials are not moved while the latter are:

(13) a. * Swiftly, John likely ran.
b. So swiftly did John likely run that he didn't even notice the fire.

Still more evidence that these adverbials are not moved comes from their behaviour with respect to VP ellipsis (Espinal 1991: 731; McCawley 1982: 96), which is strikingly similar to that of parenthetical and afterthought occurrences of adverbials, a point to which we shall be returning in the following section. As 14 shows, all three classes of adverbials behave as elements independent of the antecedent VP, as reflected in the readings indicated:

(14) a. With his X-ray vision, John located the documents and Bill did too.
= 'Bill located the documents';
≠ 'with John's/his X-ray vision, Bill located the documents'
b. John, with his X-ray vision, located the documents and Bill did too.
= 'Bill located the documents';
≠ 'Bill, with John's/his X-ray vision, located the documents'
c. John located the documents, with his X-ray vision, and Bill did too.
= 'Bill located the documents';
≠ 'Bill located the documents, with John's/his X-ray vision'

This pattern is precisely what McCawley observed for standard parenthetical expressions, as shown in 15:

(15) John talked, of course, about politics, and Mary did too.
= 'Mary talked about politics too'; ≠ 'Mary talked too'
≠ 'Mary talked, of course, about politics too'
(McCawley 1982: 96, (5a))

Consideration of VP ellipsis structures reveals a final piece of evidence that the fronted adverbials in question are not moved from positions within the VP. This is that they can acceptably apply to both conjuncts in cases like that given in 16a, even though in doing so they clearly cannot have moved from the VP:

(16) a. With their X-ray vision, John located the documents and Bill did too.
b. * John located the documents with their X-ray vision and Bill did too.

Taken together, the data given in this section provide substantial evidence that certain fronted adverbials in English with 'lower' readings are not moved from the VP, contrary to the standard view of them (e.g. Ernst 2002: 407–33). Of course, such a claim about these adverbials raises new questions about how they should be analysed, to which we turn in the next section.

3. So where in the tree are these things? Toward an 'orphan' analysis

If we accept that these adverbials are not moved, then the standard alternative is to take them to be IP or CP adjuncts (e.g. Haegeman 2002). However, on this assumption, adjuncts in this position might be expected to make a constant contribution to interpretation in having sentential scope. Yet this is clearly not the case with the adverbials in question, many of which are compatible with a range of readings, as the sentences in 17–18 indicate:

(17) Happily, John sucks lemons.
= 'It is a happy circumstance that...'
= 'John is happy to suck lemons'

(18) a. Quietly now, what is the problem? = 'Speaking quietly, tell us what...'
b. Quietly, John kept about his business. = 'John quietly kept...'
c. Very quietly for a six year-year-old, Robert crept down the stairs.
= 'Robert crept... in a quiet manner...' (Ernst 2002: 395, (8.16b))

As the paraphrases above suggest, *happily* in 17 has both sentential and subject-oriented readings; while *quietly* in 18 has sentential, subject-oriented, and manner readings. Arguably neither lexical nor structural ambiguity is an attractive option for analysing these different readings: the key is simply that the argument of the adverb is not specified.

One plausible way to account for these patterns—which has already been proposed for a range of expressions that 'bear no obvious syntactic relationship to the sentences they seem to be included in' (Espinal 1991: 726)—is to take these fronted adverbials to be 'orphans' (Haegeman 1991). Elements of this kind, as Haegeman and others have described them, occupy positions that have no hierarchical relation to the sentences with which they are associated, and *a fortiori* are not in a c-command relation to these 'host' sentences. Significantly, such a claim is consistent with the binding, island, and VP ellipsis data reviewed in the previous section. It also gives us a way to capture the observation that these adverbials are typically set off intonationally from the rest of the sentence, the idea being that speakers may signal the presence of a distinct syntactic unit by this prosodic means.

More direct evidence for an 'orphan' analysis of these adverbials is their inability to license negative polarity items when they contain polarity licensors like *only*, where the licensor needs to be in a c-command relation to the negative polarity item. This behaviour, as indicated in 19a, again contrasts with that of moved adverbials like those in sentences 19b–c (19c being taken to display focus movement), in which the negative polarity item licensors are understood to be in a c-command relation to the polarity items:

(19) a. * Only quietly, John ever got drunk.
b. Only quietly did John ever get drunk.
c. ? Only QUIETLY John ever got drunk.

An 'orphan' analysis of the fronted adverbials described here and their parenthetical and afterthought counterparts thus captures the mismatch between their position and interpretation in a manner consistent with the evidence that they are not moved and without recourse to the discontinuous structures by which McCawley (1982) sought to solve the puzzle of this mismatch.[2] What is also worth noting is that 'orphans' arguably need to be recognized as grammatical elements in any case, given the existence of a range of elements loosely associated with their host sentences (e.g. Espinal 1991) and of the 'ordinary, unembedded, words and phrases' with which 'speakers can make assertions' (Stainton 1995: 281), as illustrated in 20 and 21, respectively.

(20) a. The secretary, well-mannered as anybody, will present an apology.
b. Today's topic, ladies and gentlemen, is Nuclear Magnetic Resonance.
(Espinal 1991: 726–27, (3b), (6a))

(21) a. Nice dress
b. Black coffee with no sugar (Stainton 1995: 293, (20a), (20e))

Such considerations all point toward the existence of 'orphans' in the grammar and the plausibility of analysing fronted, parenthetical, and afterthought adverbials in 'orphan' terms.

4. Some open questions (and some tentative answers)

From the above discussion, it appears that an 'orphan' analysis can answer many puzzling questions about fronted, parenthetical, and afterthought adverbials. Admittedly, many more questions remain, adequate answers to which must await further research. It might, however, might be worth considering some of these questions—briefly, given space limitations—at least to indicate the direction that answers to them might take. Moreover, given the articulated C approach against which the discussion of 'orphans' has unfolded, it is worth asking whether an 'orphan' analysis is ultimately compatible with this approach.

Among the many questions that remain for our 'orphan' analysis is why various patterns of acceptability among supposedly 'orphan' adverbials are in fact

[2] Of course, it remains possible that parenthetical expressions take the form of adjunction structures fully integrated into the sentence, as Potts (2002) has argued for *as*-parentheticals. I discuss the difficulties of generalizing Potts's account to the adverbials discussed in this paper and certain more basic difficulties with his approach in Shaer (2003b).

consistent with a movement analysis. These patterns include the unacceptable fronting of manner adverbs over modals and of measure adverbs generally (Ernst 2002: 424) and apparent Principle C violations with fronted adverbials in configurations like those of the acceptable sentences in 10. These patterns are illustrated in 22, 23, and 24, respectively:

(22) a. * Carefully, he must ease the violin out of its case.
b. * Tightly, she must hold on to the railing. (Ernst 2002: 422, (8.94))

(23) a. * Completely, he eased the violin out of its case.
b. * Partway, Karen filled the glass. (Ernst 2002: 422, (8.95a–b))

(24) a. * In Ben's$_i$ office, he$_i$ lay on his desk.
b. * For Mary's$_i$ brother, she$_i$ was given some old clothes.
(Speas 1990: 51–52, (59b), (60b))

In fact, a plausible explanation of these patterns is available once we consider some of the properties of 'orphans' and of fronted elements more generally. By hypothesis, 'orphans' are syntactically independent of their host sentences and thus semantically integrated into their hosts only through a process of inference. We can also speculate that fronted adverbials, like other preposing devices, may serve as 'links', relating 'the information presented in the current utterance' to the 'information evoked in or inferrable from the prior context' (Birner and Ward 1994: 20–22). What this suggests is that the degraded acceptability of a sentence containing a fronted adverbial may have its source in the difficulty of inferring the relation of the adverbial not only to its host sentence but also to previous discourse. That the unacceptability of the sentences in 22–24 may have such non-syntactic sources is brought home by the acceptability of sentences with other adverbial-modal combinations, as shown in 25; by the existence of acceptable sentences, as in 10, that are structurally analogous to those in 24; and by the obvious difficulties involved in any attempt to infer the arguments of the measure adverbs in 23 (see Shaer 2003a: 249).

(25) a. Quietly, he will tell her that he loves her.
b. With the greatest of tact, she might suggest that he find suitable employment elsewhere.

The above remarks about the discourse functions of fronted 'orphans' put us in a good position to address the question of the compatibility of an 'orphan' analysis with the 'articulated C' analysis discussed above. Arguably it is impossible to find any strict incompatibility between them, since the former takes certain elements to lie beyond the sentence, while the latter concerns only elements fully integrated into the sentence. The latter can thus incorporate the results of the former without any alteration of its basic claims. Yet since the latter seeks to assimilate the patterns observed in one language to those observed in another, a significant point of tension does emerge after all. This is because of the close connection that the 'articulated C' approach posits between syntactic structure and both discourse function and illocutionary force. Of course, languages do exist—Italian and German among those that readily come to mind—for which the hypothesis of a C domain containing positions with well-defined discourse functions is a plausible one. Yet the present investigation of fronted adverbials in English suggests that such an hypothesis is far less plausible for this language, which makes very free use of 'orphans' and in which relatively few classes of adverbials can be seen as moved to LP positions. Since 'orphan' adverbials appear to fulfil the discourse functions of their moved counterparts, the tight connection between LP positions and discourse functions that is at the heart of the 'articulated C' approach and the limited degree of cross-linguistic variation that this approach envisages lead to the conclusion that its picture of the left periphery is not an accurate one.

5. Conclusion

In this paper, I have presented evidence that certain fronted adverbials in English have not moved to positions high in the tree, but are generated as 'orphans', elements syntactically distinct from their host sentences. This suggests that significant cross-linguistic differences may exist in the structure of the left periphery, and thus that reductionist approaches to this domain, like that proposed by Rizzi (1997) and others, are unlikely to shed light on its general properties. Moreover, consideration of the behaviour of 'orphans' suggests that no direct relation exists between syntactic position and either discourse function or illocutionary force, both of which are better accounted for in 'modular' terms. Finally, although the 'orphan' adverbials described here are admittedly rather strange linguistic beasts, their inclusion in the grammar offers a simple way to account for a host of refractory syntactic, semantic, and prosodic data.

References

Baltin, M. 1982. A landing site for movement rules. *Linguistic Inquiry* 13: 1–38.

Birner, Betty and Gregory Ward. 1994. *Information Status and Noncanonical Word Order in English*. Amsterdam/Philadelphia: John Benjamins.

Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.

Espinal, M. Teresa. 1991. The representation of disjunct constituents. *Language* 67: 726–62.

Haegeman, Liliane. 1991. Parenthetical adverbials: The radical orphan approach. In S. Chiba et al. (eds.). *Aspects of Modern English Linguistics: Papers Presented to Masatomo Ukaji on His 6oth Birthday*, 232–254. Tokyo: Kaitakushi.

——. 2002. Anchoring to speaker, adverbial clauses and the structure of CP. Lille: Université Charles de Gaulle-Lille III ms.

Hukari, Thomas and Robert Levine. 1995. Adjunct extraction. *Journal of Linguistics* 31: 195-226.

McCawley, James. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13: 91–106.

Newmeyer, Frederick. 2002. Interpretable features, feature-driven A-bar movement and the syntax-discourse interface. Paper presented at GLOW, Amsterdam.

Pollard, Carol & Ivan Sag. 1994. *Head-driven Phrase Structure Grammar.* Chicago/London: University of Chicago Press.

Potts, Christopher. 2002. The syntax and semantics of *as*-parentheticals. *Natural Language and Linguistic Theory* 20: 623–689.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.). *Elements of Grammar: Handbook in Generative Syntax*, 281–337. Dordrecht/Boston/London: Kluwer

——. 2002. Locality and left periphery. In Adriana Belleti (ed.). *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3. Oxford/New York: Oxford University Press.

Shaer, Benjamin. 2003a. 'Manner' adverbs and the association theory: Some problems and solutions. In Ewald Lang et al. (eds.). *Modifying Adjuncts*, 211–259. Berlin: de Gruyter.

——. 2003b. An 'orphan' analysis of long and short adjunct movement in English. Paper presented at the West Coast Conference on Formal Linguistics, University of California, San Diego, March 2003.

Shiobara, Kayono. 2003. Reexamining Condition C effects. This volume.

Speas, Margaret. 1990. *Phrase Structure in Natural Language.* Dordrecht/ Boston/London: Kluwer.

Stainton, Robert. 1995. Non-sentential assertion and semantic ellipsis. *Linguistics and Philosophy* 18: 281–96.

Taglicht, Josef. 2001. *Actually*, there's more to it than meets the eye. *English Language and Linguistics* 5: 1–16.

Zanuttini, Raffaella and Paul Portner. To appear. Exclamative clauses: At the syntax-semantics interface. *Language* 79.

Zentrum für Allgemeine Sprachwissenschaft
Jägerstr. 10/11
10117 Berlin, Germany

shaer@zas.gwz-berlin.de

Patterns of Semantic Harmonization in English: The Case of *May Well*[1]

REIJIROU SHIBASAKI
University of California, Santa Barbara

0. Introduction

The diachronic semantic-pragmatic process whereby meanings have been based in the speaker's subjective belief towards what is said, i.e. subjectification, has received considerable attention in recent years (Traugott 1989; Langacker 1990; Stein and Wright 1995; inter alia). The development of epistemic modality has been a prime target. Subsequent research has focused on either general or cross-linguistic properties or evolution of modal meanings (Bybee et al. 1994). Recently, one approach appears to gain prominence in studies of language change. It is statistical studies that can be a valuable tool in providing relatively firm evidence for language change. This approach is not new but rather a revival of an older notion of 'frequency' to give an account of the evolution of language (see Martinet 1964: ch.6). Reasonable though this approach seems to be, most preceding studies in grammaticalization and language change in general have been theory-oriented with no reliable amount of examples, just providing general pathways of change. It is true that there emerge a subset of works building on the notion of frequency, synchronically or diachronically, such as Bybee (2001) and Krug (2000). But, it is still necessary to more develop statistical parameters common among many languages for better understanding degrees of morpho-syntactic and semantic changes. Hopper and Traugott (1993: 112) argue under that "[T]here is an urgent need for additional reliable statistical studies of a variety of phenomena in which early grammaticalization appears to be involved". Frequency or quantitative approaches to language change are time-consuming but inevitable to reinterpret language change in a down-to-earth way.

Based on a text-based quantitative approach, I will illustrate the development of epistemic meanings in the modal verb-adverb construction in the history of English,

[1] I am grateful to Mary Bucholtz, Patricia Clancy, Sandra A. Thompson, Makiko Takekuro and Shin'ichi Nishikawa for their comments that have contributed directly to this paper. I am also grateful to Fumie Nakamura for her invaluable support. Comments from John W. Du Bois at the conference are here gratefully acknowledged. The remaining faults are of course all my own.

with special focus on *may well*. The arguments from this study are that: 1) only semantically cohesive co-occurrences have been selected and conventionalized over history, i.e. semantic harmonization (to be explained in section 1); 2) *may well* is one of the general phenomena of semantic harmonization, and; 3) *may well* is an idiomatic expression of the modal verb-adverb construction. While most preceding research has uncovered the semantic expansion of modals themselves, I assume that a frequency-based constructional approach is a useful device for better understanding language change.

1. Theoretical Implications

There are several research projects that focus more or less on the modal verb-adverb co-occurrence in several languages from a synchronic perspective (for English see Jacobson 1964; 1975; Lyons 1977; Hoye 1997; for French see Guimier 1996) and from a diachronic perspective (Jacobson 1981; Shibasaki 1998; Traugott and Dasher 2001: 3.4.1; all about English). For example, Lyons (1977) discusses the semantic cohesiveness between modal verbs and adverbs. He concludes that they normally reinforce each other in meaning in a modally harmonic way. Let us consider the following pairs (*ibid.*: 807-8).

(1) a. He may possibly have forgotten.
b. He may certainly have forgotten.

In (1a), the semantic relation between the modal verb and the adverb is within the same scope of likelihood; therefore, they can make a single modality that is incorporated in the clause. In (1b), however, the modal verb *may* cannot be fused semantically or pragmatically with the adverb *certainly*, because each lexical item resides in the different level of likelihood; therefore, the interpretation is ambiguous. The important thing is how much epistemic scales of a modal verb and an adverb can be coalesced into the same scope of modality. If a modal verb and an adverb are semantically within the same scope of likelihood, they tend to co-occur harmonized in meaning. I call this phenomenon semantic harmonization.

Hoye (1997: 240) sets forth scales of likelihood based on his corpus-based analysis, which are illustrated in Table 1 (with slight modification). One of his important arguments is that "the unmodified modal conveys possibility, whereas the combination designates probability" (*ibid.*: 234; also see 241); in other words, when a modal verb co-occurs with an adverb, the scale of likelihood becomes higher. It is also suggested that not all combinations of modal verbs and adverbs are possible judging from his corpus-based analysis, but a modal verb and an adverb from the same or closer scale of likelihood have a strong tendency to co-occur. This finding gives support for Lyons' comments on (1) and our view of semantic harmonization.

May well has been considered as being the most problematic collocation; however, most researchers seem to have reached a consensus that *may well* is an idiomatic expressions. Hoye (1997: 235) assumes that "MAY WELL appears to be a modal idiom, on a par with other verbal constructions like HAD BETTER or WOULD RATHER which are widely accorded the status" (see also Palmer 1995: 210). Furthermore, *may well* shows some rigid constructional aspects, as Hoye (1997: 234) argues that "MAY WELL is invariable and resistant to any form of modification; paradigmatic substitution of other adverbs, while possible, disrupt the integrity of the meaning this particular collocation conveys". The following analysis gives support for the view that *may well* is an idiomatic expression of the modal verb-adverb construction in terms of semantic harmonization and frequency.

Table 1: Epistemic modal and modal adverb classification (Hoye 1997: 240)

A *Possibility*	B *Probability*	C *Certainty*
MIGHT MAY COULD CAN	SHOULD OUGHT TO WOULD WILL	MUST CAN'T
POSSIBLY CONCEIVABLY PERHAPS MAYBE	PROBABLY QUITE LIKELY MOST LIKELY WELL	CERTAINLY DEFINITELY INDEED PRESUMABLY SURELY FOR CERTAIN OF COURSE UNDOUBTEDLY NECESSARILY

2. Data

As is often pointed out, expressions scribed in Middle English texts are considered to reflect colloquial expressions of those times. For example, Burrow and Turville-Petre (1996: 5) assume that "...authors in the twelfth, thirteenth and fourteenth centuries generally wrote the English that they spoke – whether in London, Hereford, Peterborough, or York". Suppose that a grammatical form is more or less a reflection of its colloquial expression of those times, then it would be possible to comparatively investigate morpho-syntactic forms between carefully selected Middle English texts, those in Early Modern English dramas, and those in later stages including Present Day English conversation. This view makes it implausible to draw the line between semantics and pragmatics because meanings of each expression are rooted in any immediate discourse (Traugott and Heine 1991; see Jacobs and Jucker 1995 for two subtypes of historical pragmatics). That is to say, this study examines semantic harmonization of the modal verb-adverb construction in colloquial usage in the history of English.

In order to examine colloquial usage at each synchronic stage, I chose the following verse texts. I mainly analyzed the texts illustrated in (2). The frequencies of each modal verb-adverb construction at each stage are to be illustrated later.

(2) Primary

Old English (=OE):	University of Virginia Electric Text Center
Middle English (=ME):	*The Canterbury Tales* (=*CT*, 14C)
Early Modern English (=EModE):	Shakespeare's Works (16/17C)
Present Day English (=PDE):	Brown/London-Lund Corpora (20C)

(3) Supplements

13C: *The Owl and the Nightingale* (=*O&N*, verse)
14C: *Sir Gawain and the Green Knight* (verse)
15C: *Mankind, The Wakefield Pageants* (early dramas)
16C: *Mundus et Infans* (an early drama)

3. Analysis

Goossens (1982: 78) argues that clear epistemic meanings are hard to find out in ME. Bybee (1988: 258-9) makes a similar observation: "[A]bout one-third of the examples of *may* in *Sir Gawain and the Green Knight* can be interpreted as either root or epistemic possibility; the rest are unambiguously root". I agree to this point; however, I assume that it becomes possible to present epistemic meanings in outline from the viewpoint of semantic harmonization of the modal verb-adverb construction. The following example shows a semantic contrast between the co-occurrence of *may well* and a single occurrence of *may*. Examples in focus are underlined from here on.[2]

(4) Arthur: *Neverþelece to my mete I <u>may</u> me <u>wel</u>*
nevertheless to my food I may:PRES.1sg myself well
dres for I haf sen a selly, I <u>may</u> not forsake
get.ready because I have seen a wonder I may not refuse
'Nevertheless, I am probably allowed to begin my meal, because I have seen an extraordinary sight (and) I can not deny (it).'
(c. 1375-1400 *Gawain* 474-5)

In the utterance of (4), the speaker Arthur is conceding: 'this is not entirely the expected Christmas fare' (Burrow and Turville-Petre 1996: 198). That is to say, the speaker Arthur would have begun his meal if his expectation had been met, but it was cancelled (i.e. not met). Therefore, the speaker's speculation on the current situation brings about an epistemic meaning by using the *may well* construction (cf. Traugott and Dasher 2001: 21ff and passim for the notion of generalized invited inference, GIIN). Permission reading would also be fine. In my opinion, *may well* has already established its epistemic possibility meaning at least in Late ME, and if

[2] The glossing conventions are as follows: INF=infinitive; PRES=present tense; PST=past tense; sg=singular; 1=first person; 3=third person.

speakers of those days want to distinguish the original root meaning from the epistemic meaning, they tend to collocate *may* with *well*, as in (4).

This semantic harmonization began to emerge as early as OE. In (5), *may well* indicates epistemic possibility and it is part of a quoted speech.

(5) *Cweð he: wel þet swa mæg*
say:PST.3sg he well that so may:PRES.sg
'He said: that will probably (be) so.' (c. 900 tr. *Bæda's Hist.* II.i 110)

In ME onward, *may well* continues to indicate or rather reinforce epistemic possibility in various contexts. In (6), the writer makes a speculation that the nightingale is most likely to be thoughtful judging from the context, while in (7), the negative conditional clause *but it be a fool* invites an inferential interpretation from *may well* as being 'it is no wonder/with good reason'.

(6) *þe nyhtegale fat and fyhte & hauhful waf*
the nightingale sit:PST.sg and fight.INF and thoughtful was
& wel myhte for þe vle fo ifpeke hadde and
and well might because the owl so spoken had and
hire fpeche fo iladde.
her law-suit so led
'The nightingale sat and fought, and was thoughtful and might probably be (so), because the owl had spoken and led her case in such a way.'
(a. 1250 *O&N* Jes. Oxf. 1291-4)

(7) *wel may men knowe, but it be a fool,*
that every part dirryveth fromhis hool
'It is no wonder, unless they are not fools, that men know that every part derived from the whole.' (c. 1388-1400 *CT.*, A Kn 3005-6)

In Late ME, *may well* is used in the impersonal construction. There are several examples of *may well* that appear in the impersonal construction in *The Canterbury Tales*, yet here I will illustrate one example from the following text, as in (8). There might be much earlier examples of this kind if we refer to a larger corpus; however, it is good to know that such expressions begin to emerge around this stage.

(8) Cayphas: *so may I well seme, myself if I say it*
'It will no wonder become me, if I myself say it.'
(c. 1400-450 *Coliphizacio* 289-90; in Cawley ed.)

The verb *seme* 'to become' is usually an impersonal verb in this sense, so the semantic subject is not the grammatical subject *I*. Therefore, the speaker's inference is projected towards the proposition, as reflected in the translation (cf. Cawley 1958: 121, note 260). The conditional clause also reinforces the inferential reading, i.e. epistemic possibility, in this context.

In (9), the character Nowadays becomes scared that another character Mischief (Myscheff) would chop off his head in order to kill the pain there. Considering his own cowardice, Nowadays thinks of it no wonder that he would be called fool. It is hard to diagnose degrees of epistemic meanings because contextual cues are limited. In this case, scales of epistemicity seem to be the same as earlier examples.

(9) Nowadays: *Ȝe, Cristys crose, wyll Ȝe smyght my hede wey?*
Ther wer on and on! Oute! Ȝe xall not assay.
I myght well be callyde a foppe.
'Yea, Christ's cross! Will you smite my head away? There were one after another! (or There! Where? On and On!) Out! You shall not assay – No wonder I might be called a fool.'
(c. 1465-70 *Mankind* 442-4; in Eccles ed.)

In the following example, *may well* may possibly be considered to indicate more inferential or subjective meaning than other examples. Let me explain about it.

(10) Hermia: *Lie further off, in human modesty;*
Such separation as may well be said
Becomes a virtuous bachelor and a maid
(c. 1595-96 *A Mid-Summer Night's Dream* II ii 56-8)

In this context, the speaker Hermia assumes that such separation, i.e. not to share a bed, is obligatory for a not-yet-married couple judging from the current belief of those days. Therefore, the speaker sets forth a relatively strong argument on their behaviors. In other examples earlier than (10), it was hard to find out contextual cues as strong as in (10). I assume then that in (10), the scale of epistemicity is stronger than other examples and the speaker's choice of *may well* (not *may*) fits this context, supporting her strong argument.

The following example shows a semantic contrast between the co-occurrence of *might well* and a single occurrence of *might*.

(11) Antonio: *Th'offence is not of such a bloody nature,*
Albeit the quality of the time and quarrel
Might well have given us bloody argument.
It might have since been answered in repaying
What we took from them, which for traffic's sake
Most of our city did. (c. 1601 *Twelfth Night* III iii 30-5)

By using *might well*, the speaker Antonio assumes that such a severe fight, i.e. killing, would have happened considering the bad relation between Antonio and the city, although the fight was not so severe in reality. This context invokes a quite strong inference towards the possibility of such a fight with which *might well* seems to harmonize better than *might* alone. In the context of a single occurrence of *might*, the inference appears to be less strong. The speaker simply thinks that repaying what they did to the city would be enough for their reconciliation; however, his idea is just a weak speculation. It seems to be impossible to exchange *might well* with *might* in this context. From examining these semantic contrasts in discourse, it can be said that the combination of *may well* has a tendency to imply a stronger epistemicity than a single occurrence of *may* does.

A much stronger epistemic coloring of *may well*, which is close to probability, is hard to find out in my texts around this stage. In PDE, however, this semantically harmonized expression *may well* seems to have semanticized probability in it, as in (12).

(12) *That... <u>may well</u> be over the line.* (*U.S. News*, Aug. 10, 1998)

In this example, *well* modifies the preceding modal *may* to strengthen the likelihood of the topic event with certain evidence (in this case, it implies that the suspicious scandal of the former President Bill Clinton with Monica Lewinsky). While it is possible to simply use *may* in this context, it seems apparent that the article writer emphasizes or rather concludes by using *may well* that the suspicious scandal happened at all.

4. Negation

In this section, I will briefly account for the semantics of *may well* with negative particles. Bybee et al. (1994) suggests that while admitting there are many examples of semantically indeterminate use of *may* in *Sir Gawain and the Green Knight*, the meaning of *may* can be regarded as the root possibility reading when *may* co-occurs with the negative particle. The following is excerpted from Bybee et al. (1994: 198).[3]

(13) *For mon may hyden his harmes, bot vnhap ne may hit.*
'For a man may/can hide misfortunes, but he cannot undo them.'

[3] The original text of *Sir Gawain and the Green Knight* seems not to be specified in Bybee et al. (1994). However, judging from the example cited, it seems that the source would be the second edition of the text revised by Norman Davis, which was originally edited by J. R. R. Tolkien and E. V. Gordon, from Oxford University Press. The following is from another text (Burrow 1972):

(i) *For none may hyden his harme bot unhap ne may hitte*
'For a man may hide his (spiritual) harm, but cannot unfasten it (get rid of it).'
(or 'For no-one can conceal his guilt without misfortune befalling.')

This line can be interpreted in two ways. When *unhap* and *hitte* are interpreted as a verb 'unfasten' and a pronoun 'it' respectively, the translation corresponds to the first. In this case, *non* is emended as *mon* 'one'; when *unhap* and *hitte* are interpreted as a noun 'misfortune' and a verb 'hit' respectively, the translation corresponds to the second (*ibid.*: 124).

(c. 1375-1400 *Gawain* 2511)

I totally agree to their suggestion, since examples of *may* from my texts around this stage only have the root possibility reading when they co-occurs with the negative particles. Important is that the meaning of *may well* also indicates root possibility when it co-occurs with negative particles, as follows.

(14) Vxor eius: *Go to another stede! I may not well queasse;*
Ich fote that ye trede goys thorow my nese
So hee.
'Go somewhere else! I cannot well breathe. Every step that you tread goes through my nose so loudly.'
(c. 1400-450 *Secunda Pastorum* 487-9; in Cawley ed.)

Remember that *may well* is considered to be an epistemic modal idiom resistant to any modification (Hoye 1997: 234). This diachronic frequency-based study also suggests that *may well* has gradually fixed its idiomatic formation, because it has come to disfavor even being negated over history. Small though they are, my texts show that five examples of *may well* with the negative particle in *The Canterbury Tales*, two in Shakespeare's works, and zero in Brown and London-Lund corpora. This observation seems to point out that *may well* has become semantically or syntactically a fixed expression. Further research will reveal more about this diachronic path.

5. Summary and Discussion

In this section, I will summarize the findings thus far. Tables 2 and 3 illustrate patterns of semantic harmonization from a diachronic perspective with their frequencies. Among several significant findings about the evolution and selection of the modal verb-adverb construction, I will focus on two points.

First, as in Table 2, both the number and the proportional frequencies of *may well* have been increasing, while other combinations of modal verbs with *well* have been decreasing over history. I will argue then that the longevity of this co-occurrence can be attributed to the semantic cohesiveness of *may* and *well*, which is fused into the same scope of modality; namely, semantic harmonization. The relatively high frequency also suggests that *may well* has been conventionalized more or less as a fixed epistemic expression over history; namely, a modal idiom. Building on his corpus- based analysis of PDE, Hoye (1997: 234) points out that "MAY WELL is invariable and resistant to any form of modification; paradigmatic substitution of other adverbs, while possible, disrupts the integrity of the meaning this particular collocation conveys". My texts show that almost all examples of *may well* appear unmodified even with any other adverbs in OE through PDE, except some intensifying adverbs like *very*. Moreover, *may well* has come to disfavor even being negated over time, as I explained in section 4. *May well* is often regarded as being the most problematic among the modal verb-adverb construction, since it represents a unity of meaning which is not retrievable from the sum of its individual parts (*ibid.*: 227, note 3; 232ff). However, this frequency-based diachronic study draws a relatively firm conclusion that *may well* is one of the general phenomena of semantic harmonization, as shown in Tables 2 and 3. Furthermore, despite the emergence of many (modal) adverbs in the fourteenth century onward, it is the adverb *well* that *may* has come to choose to form the epistemic modal idiom, as in Table 3. This means that *may* and *well* are most harmonized in its combination synchronically and diachronically, and this study evidences it with their frequencies. It can be considered then that *may* and *well* have developed as a modal idiom.

Table 2: Co-occurrences with *Well*

	***The Canterbury Tales* (14C)**			***Shakespeare*'s Works (16/17C)**			**Brown/London-Lund Corpora (20C)**		
	PST	PRES	Total	PST	PRES	Total	PST	PRES	Total
May well	5	42	47	17	30	47	41	32	73
Can well	22	13	35	12	9	21	2	8	10
Dare well	0	17	17	...	1	1	...	...	...
Ought well	13	0	13	...	...	...	...	...	...
Shall well	2	9	11	1	5	6	...	...	...
Will well[4]	2	1	3	...	...	...	...	...	...

Table 3: Adverbs Frequently Co-occurring with *May*

	Old English Verse	***The Canterbury Tales***	***Shakespeare*'s Works**	**Brown/ London-Lund Corpora**
May well	5	47	47	73
May as well	...	2	16	22
May easily	4	6	8	2
May lightly	...	6	...	...
May chance	...	...	6	...
May haply	...	...	5	...
May happily	...	...	5	...
May perhaps	...	...	3	8
May possibly	...	...	...	4
May hardly	...	...	2	...
May reasonably	...	1	1	...

The other is concerning the gradual development of epistemic meanings in *may well*. The context where *may well* appears in my texts involves certain information to give support for the diagnosis of the relatively strong degrees of epistemicity as

[4] Interestingly, it is explained in the glossary of *Havelok the Dane* (Smithers 1987) that *wile well* 'will well' is a collocational expression in those days, while either *may well* or *can well* are not, although the latter two appear more frequently than the former.

early as OE. Degrees of epistemicity are hard to measure out; however, it is possible to do it if we examine the semantic contrast between the co-occurrence of *may well* and a single occurrence of *may* at a same synchronic stage, as in (11). In almost all cases in my texts, the epistemicity of *may well* is stronger than that of *may*. Hoye's (1997: 98) comments are compatible with this idea: "MIGHT and MAY favour adverbs which indicate the speaker's degree of confidence". Meaning is quite hard to specify even in context. But several contextual cues make it possible to diagnose degrees of epistemic meanings.

6. Concluding Remarks

In this study, I have given a broad outline of the diachronic status and transition of the modal verb-adverb construction, with special focus on *may well*. *May well* has been regarded as being the most problematic in this construction, and then regarded as a modal idiom. However, once we pay attention to the relative frequencies of *may well* in comparison of others, it becomes apparent that *may well* is the most frequently used expression in the modal verb-adverb construction, synchronically and diachronically. Furthermore, the relatively high frequency and longevity of *may well* can be attributed to the semantic cohesiveness of *may* and *well* coalesced into the same scope of modality; namely, semantic harmonization. It is true that this small study is just a preliminary stage for a fuller diachronic analysis of *may well* and the modal verb-adverb construction in general. But findings from this diachronic study meet or rather reinforce those from a synchronic study by Hoye (1997). It is hoped then that a combination of diachronic and synchronic frequency-based studies will further uncover what is yet to be uncovered in English and other languages.

References

(1) Primary Sources

Benson, Larry D. (ed.) 1987. *The Riverside Chaucer*. Boston: Houghton Mifflin.

Burrow, J. A. (ed.) 1972. *Sir Gawain and the Green Knight*. Middlesex, England: Harmondsworth.

Crawley, C. A. (ed.) 1975. *The Wakefield Pageants*. New York: Manchester UP.

Eccles, Mark. (ed.) 1969. *The Macro Plays. EETS 262*. London and New York: Oxford UP.

Farmer, J. S. (ed.) 1909. *Mundus et Infant* (facsimile), London and Edinburgh.

Grattan, J. H. G. and G. F. H. Sykes. (eds.) 1935. *The Owl and the Nightingale, EETS extra series 119*. London and New York: Oxford UP.

Oizumi, Akio. (ed.) 1991-92. *A complete concordance to the works of Geoffrey Chaucer*. 10 vols. Hildesheim and New York: Olms-Weidmann.

Smithers, G. V. (ed.) 1987. *Havelok the Dane*. Oxford: Clarendon Press.

Spevack, Marvin. (ed.) 1969. *A complete and systematic concordance to the works of Shakespeare*. Hildesheim: George Olms Verlagsbuchhandlung.

Tolkien, J. R. R. (ed.) 1964. *Sir Gawain and the Green Knight*. Oxford: Oxford UP.

University of Virginia Electric Text Center, The Complete Corpus of Old English. (http://etext.lib.virginia.edu/oec.html)

(2) **Secondary Sources (selected)**

Burrow, J. A. and T. Turville-Petre. 1996. *A book of Middle English*. 2nd ed. Oxford: Blackwell.

Bybee, Joan L. 1988. Semantic substance vs. contrast in the development of grammatical meaning. *Proceedings of the fourteenth annual meeting of Berkeley Linguistics Society*: 247-64.

Bybee, Joan L. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan L., William Pagliuca and Revere D. Perkins. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Goossens, Louis. 1982. On the development of the modals and of the epistemic function in English. In Anders Ahlqvist (ed.) *Papers from the fifth ICHL*, 74-84. Amsterdam: John Benjamins.

Guimier, Claude. 1996. *Les adverbes du français*. Paris: Ophrys.

Hopper, Paul and Elizabeth C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Hoye, Leo. 1997. *Adverbs and modality in English*. London: Longman.

Jacobs, Andreas and Andreas Jucker. 1995 The historical perspective in pragmatics. In Andreas Jucker (ed.) *Historical pragmatics*, 3-33. Amsterdam: John Benjamins.

Jacobson, Sven. 1964. *Adverbial positions in English*. Stockholm: Studentbok.

Jacobson, Sven. 1975. *Factors influencing the placement of English adverbials in relation to auxiliaries*. Stockholm: Almquvist & Wiksell International.

Jacobson, Sven. 1981. *Preverbal adverbs and auxiliaries*. Stockholm: Almquvist & Wiksell International.

Krug, Manfred G. 2000. *Emerging English modals*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. Subjectification. *Cognitive Linguistics* 1: 5-38.

Lyons, John. 1977. *Semantics, vol.2*. Cambridge: Cambridge UP.

Martinet, André. 1964. *Elements of general linguistics*. Trans. by Elisabeth Palmer. Chicago: University of Chicago Press.

Palmer, F. R. 1995. *Modality and the English modals*. 2nd ed. London: Longman.

Powel, M. 1992. The systematic development of correlated interpersonal and metalinguistic uses in stance adverbs. *Cognitive Linguistics* 3: 75-110.

Shibasaki, Reijirou. 1998. Historical analysis of modal verb-adverb co-occurrences in English. *Kansai Linguistics Society* 18:186-196.

Stein, Dieter and Susan Wright (eds.) 1995. *Subjectivity and subjectification*. Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 1989. On the rise of epistemic meaning: An example of subjectification in semantic change. *Language* 65: 31-55.

Traugott, Elizabeth C. and Bernd Heine. 1991. The semantics-pragmatics of grammaticalization revisited. In Elizabeth C. Traugott and Bernd Heine (ed.), *Approaches to grammaticalization* vol.1, 189-218. Amsterdam: John Benjamins.

Traugott, Elizabeth C. and Richard Dasher. 2001. *Regularity in semantic change*. Cambridge: Cambridge UP.

Reijirou Shibasaki
Department of Linguistics
University of California, Santa Barbara
Santa Barbara, California 93106-3100
U.S.A.

reijirou@umail.ucsb.edu

Constraints on Synaesthesia

YOSHIKATA SHIBUYA* and HAJIME NOZAWA∞
*University of Manchester, ∞Kyoto University

0. Introduction

Synaesthesia is the phenomenon that involves mapping from one sense modality onto another.[1] For example, *warm color* is a case of synaesthesia, since it involves the mapping from the tactile sense onto the visual sense. This paper introduces a cognitive model called the "Physiological = Psychological" model and discusses constraints on synaesthesia.

1. Synaesthesia and Cross-linguistic Survey

It is well known that synaesthesia shows a certain directionality of possible mappings (e.g., Williams 1976; Yamanashi 1988). Williams (1976: 464), for example, writes as follows:

- If a touch-word transfers, it may transfer to taste (*sharp tastes*), to color (*dull colors*), or to sound (*soft sounds*).
- Taste-words do not transfer back to tactile experience or forward to dimension or color, but only to smell (*sour smells*) and sounds (*dulcet music*).
- There are no primary olfactory words in English (i.e., none historically originating in the area) that have shifted to other senses.
- Dimension lexemes transfer to color (*flat color*) or to sound (*deep sounds*).[2]
- Color-words may shift only to sound (*bright sounds*).
- Sound-words may transfer only to color (*quiet colors*).

Figure 1 shows Williams' generalization. It shows that it is possible for the tactile

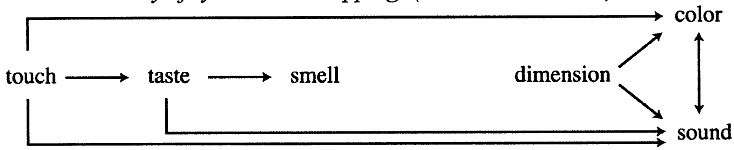
---

[1] We are grateful to William Croft and Masa-aki Yamanashi for their helpful comments on an earlier draft of this paper. We would also like to thank our native informants Chiarung Lu, Che-ho Lee, Efrosini Deligianni, Helen Bissett, and Willem Hollmann, for their patience and help.

[2] In this paper, we do not discuss dimensional adjectives.

and/or gustatory senses to be mapped onto the other senses, but not in the opposite direction.

Figure1. Directionality of synaesthetic mappings (Williams 1976: 463)



Williams suggests that his generalization is also applicable to explain synaesthesia in other languages. According to his generalization, for example, mapping from the tactile sense onto the visual sense (color) is possible, but not the other way around. On first sight, his generalization might seem to be valid, since, for instance, while those conforming to the possible mapping are found acceptable cross-linguistically (e.g., *re4de yian2se4* [Chinese: 1], *warme kleur* [Dutch: 1], *warm color* [English: 1], *θερμό χρώμα* [Greek: 1], *atatakai iro* [Japanese: 1], *ttatteushan saeg* [Korean: 1]), those conforming to a mapping assumed to be impossible by his generalization are also likely to be found anomalous (e.g., *hong2de wen1du4* [Chinese: 0], *rode temperatuur* [Dutch: -1], *red temperature* [English: -1], *κόκκινη θερμοκρασία* [Greek: -1], *akai ondo* [Japanese: -1], *ppalgan ondo* [Korean: -1]).[3]

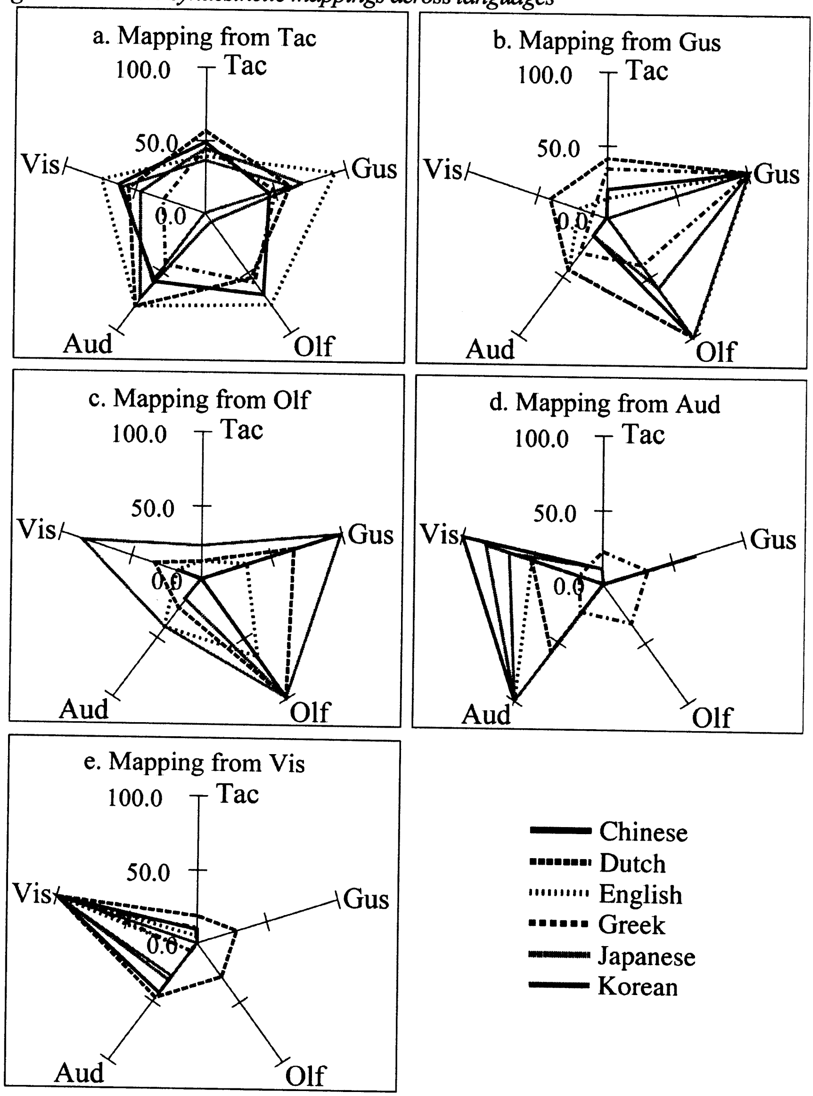
We conducted a cross-linguistic survey. Through careful observation of the results, we will show that Williams' generalization suffers from several empirical anomalies. Our target languages were Chinese, Dutch, English, Greek, Japanese, and Korean. We chose a variety of adjectives and nouns for each sense modality. Words of the tactile sense in English, for example, included the adjectives *soft* and *warm*, and the nouns *touch* and *pain*. The visual sense, for example, included adjectives such as *dark* and *red* and nouns such as *color* and *pattern*. We combined all the adjectives in each sense modality with the nouns in each sense modality. For example, for the adjective *warm*, we made phrases such as *warm touch*, *warm pain*, *warm color*, *warm pattern*, and so forth by combining the adjective with other nouns in all the sense modalities, and then examined the acceptability of the phrases.

In each language, we translated the chosen adjectives and nouns so that they corresponded with those in the other languages. The total number of phrases was 341 in each language (31 adjectives and 11 nouns). We divided the number of possible

[3] In this paper, the level of acceptability is shown by three grades on the basis of the judgments made by our native consultants: 1 indicates high acceptability, 0 acceptable if given a particular context, and -1 very low in acceptability.

combinations by the total number of the phrases in each category. Figure 2 shows the percentage of possible mappings in each sense modality. Examples which are counted as acceptable in order to make these graphs are those judged by the subjects as 1 and 0 on our acceptability grading scale.

Figure 2. Possible synaesthetic mappings across languages



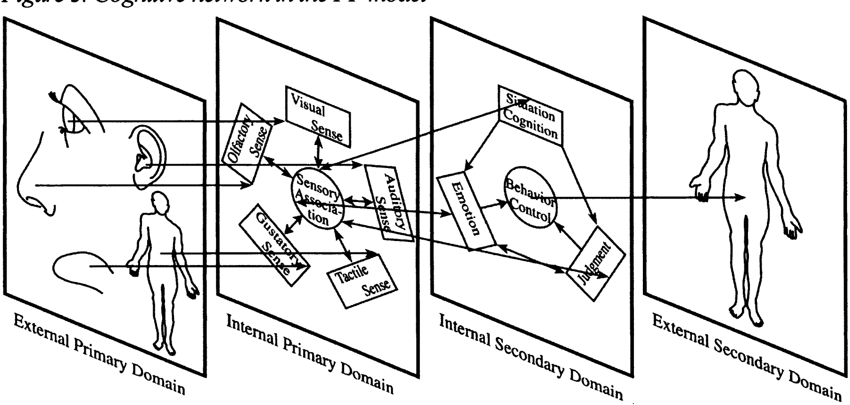
As the graphs show, despite some idiosyncrasies, a certain directionality of possible mappings is observed across the languages. Figure 2 (a) shows that mapping from the tactile sense onto the others is commonly possible. Figure 2 (b) shows that the possibility of mapping from the gustatory sense to the tactile, visual, and auditory senses is low. Figure 2 (c) shows that the possibility of mapping from the olfactory sense to the tactile, visual (except Korean) and auditory senses is low. Figure 2 (d) shows that mapping from the auditory sense to the tactile, gustatory (except Japanese) and olfactory senses is likely to be very low across the languages. Figure 2 (e) shows that mapping from the visual sense to the tactile, gustatory and olfactory senses is very low, except for Dutch.

The target languages consist of both Indo-European and non-IE languages. Therefore, one cannot ascribe such universal patterns to the historical relationships of these languages. In this paper, we argue that synaesthesia should be explained in physiological terms. We will show that our accounts using the Physiological = Psychological model provide a systematic explanation of the universal patterns, as well as the idiosyncratic mappings as shown in the graphs.

2. The Physiological = Psychological Model

The Physiological = Psychological (PP) model is a cognitive model based on the findings of physiological and anatomical studies in brain science. In this model, psychological and cognitive phenomena are equated with physiological activities of the brain. It is due to this equation that the PP model argues that meanings are reactions in the brain against linguistic stimulus (e.g., linguistic sound or letter) and a process of language comprehension is a series of brain reactions.

Figure 3. Cognitive network in the PP model



The PP model designs types of cognitive domains and their interrelations according to the anatomical structure of the brain. The present version of the model assumes four major domains: the External Primary Domain, the Internal Primary Domain, the Internal Secondary Domain, and the External Secondary Domain. Each of these includes various sub-domains.[4] The process of comprehension is described as a variety of activities in these cognitive domains. Figure 3 is a sketch of the cognitive network in the PP model.

We describe the meaning of the word *fire*, for example, as in Figure 4. The figure shows the relationship between the meaning of *fire* and the bodily experience of fire. We experience fire in daily life and the brain learns what types of features are perceived with it. The experience of fire includes warmth, a burning smell, shape and colors of a flame, sound of sparks, etc. In our accounts, a set of reactions in the Tactile, Auditory, Visual, and Olfactory Senses take place through the Sensory Association.[5] The meaning of *fire* is a compound of various reactions in the brain.

Figure 4. Meaning of fire

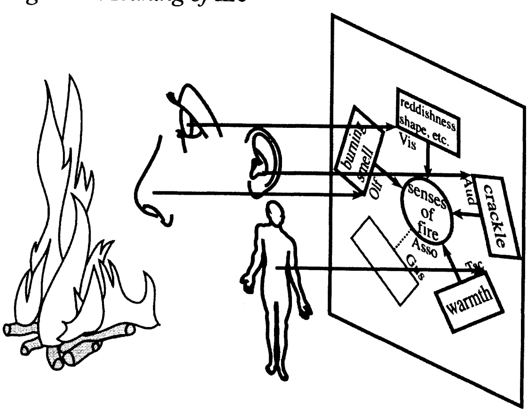


Figure 4 is not the description of the "process of comprehension" of *fire*. In the PP model, the process of linguistic comprehension is equated with a series of reactions in the various cognitive domains that are in associative relation with given linguistic stimulus. (1) is a schematic description of the reactions assumed to occur following the input of *fire*.[6]

[4] In this paper, due to the space, we limit our accounts of the structure of the PP model to a minimum. Interested readers are directed to Shibuya, Nozawa, and Kanamaru (in press) for a more detailed illustration of this new cognitive model.

[5] Note that the first letter of each sense modality here is written in capitals. This is how we describe a process in the sub-domains in the PP model.

[6] We distinguish the description of a process from the description of mental content. In the description of

(1) *Situation*:

Occurrence of the linguistic sounds of the word *fire*

Reactions in the hearer:

- Vibration of the eardrum (tympanum) [the linguistic sounds of *fire*]
- Reaction in the auditory nerve (inner ear nerve) [the linguistic sounds of *fire*]
- Reaction of the Auditory Sense [the linguistic sounds of *fire*]
- Reaction of the Sensory Association [the association between the auditory sense of the linguistic sounds of *fire* and the senses related to the elements of fire]
- A series of reactions in the sensory areas (caused through the reaction of the Sensory Association following the linguistic input of *fire*)

3. Constraint of Sensory Experiences

A particular sense tends to co-occur with other certain senses in daily life. In the case of experiencing warm objects, for instance, one learns the tactile sense of warmth and other sensory experiences such as the visual sense of reddishness of the objects as associated co-occurring senses. As illustrated above, the experience of fire includes the tactile sense of warmth as well as other senses such as the color of fire (reddishness), burning smell, etc. Likewise, the experience of electric stoves includes warmth and the colors of the heater (reddishness), as well as the shape and electric noises, etc.

Learning of simultaneous proximity between two senses enables the activation of one sense to cause the co-activation of the other sense. In the case of warm objects, the simultaneous proximity of the tactile sense of warmth and the visual sense of reddishness enables the activation of the former to cause the reaction of the latter. We argue that synaesthesia is processed through activation of co-occurring senses. (2) is a schematic description of how one processes *warm color*. The "mismatch" of sense modalities of tactility and vision is solved by the co-occurring associative relation between them, through a variety of reactions following the linguistic inputs of *warm* and *color*.

(2) *Situation*:

Occurrence of the linguistic sounds of *warm color*

Reactions in the hearer:

- Reaction of the Auditory Sense [the linguistic sounds of *warm*]
- Reaction of the Sensory Association [the association between the auditory sense of the linguistic sounds of *warm* and the tactile sense of warmth]
- Reaction of the Tactile Sense [warmth]

---

question and [Y] as the "reaction corresponding to Y in X" (mental content).

- Priming in the Sensory Association [the association between the tactile sense of warmth and other senses that are in strong associative relation with warmth]
- Reaction of the Auditory Sense [the linguistic sounds of *color*]
- Reaction of the Sensory Association [the association between the auditory sense of the linguistic sounds of *color* and the visual sense of some colors]
- Reaction of the Visual Sense [reddishness]

We argue that comprehension of *warm color* is based on the sensory co-occurrence of different senses which one learns with warm objects. Experience of warm objects is of course not limited to speakers of English, but shared by speakers of other languages. In our accounts, the cross-linguistic availability of similar synaesthetic expressions stems from the similar bodily experiences one gets with an object and the physiological function of the brain that learns the sensory co-occurrence of different senses.

Sensory co-occurrence explains how one comprehends a given synaesthetic expression, but it does not account for the directionality of possible mappings. In order to explain such a problem, we need to discuss the functional differences of sense modalities.

Human beings have developed various types of sense modalities in the course of evolution. Each sense modality differs in its properties. For example, we can perceive an object through the olfactory, visual, and auditory senses without making any bodily contact, but the other senses such as the tactile and gustatory senses are different in this respect, since perception by these senses involves bodily contact with the object.

Having various types of senses with different functions is crucial for adaptive purposes, since it helps us to avoid risking our lives. For us, for example, the best way to examine the property of an unknown object is to use the visual or olfactory senses, because otherwise if the object turns out to be deadly poisonous, it is too late if we have touched or tasted it.

We argue that our adaptive behavior with the use of sense modalities fundamentally forms what we call the "structure of sensory experiences". At the fundamental level, functional differences in sense modalities affect the frequency of sensory co-occurrence (i.e., the strength of sensory association between different sense modalities), and frequency of sensory co-occurrence is responsible for the formation of such a structure.

In terms of the tactile and visual senses, for example, the structure of sensory experiences is that the sensory association from the tactile sense to the visual sense is stronger than that from the visual sense to the tactile sense. This is due to the frequency of sensory co-occurrence that takes place in bodily experiences. Use of the visual sense does not involve bodily contact, but that of the tactile sense is likely to involve the co-use of the visual sense. For example, touching a warm object normally also involves the visual information of the object, but looking at a warm object only provides the visual information.

Synaesthesia as discussed in this paper consists of a modifier and a modified element. In the semantic structure of synaesthesia, the modifier should evoke the sense which is in a strong associative relation with the sense evoked by the modified element. As mentioned, *warm color* and its equivalent expressions are cross-linguistically acceptable, whereas *red temperature* and its equivalent expressions are not. Given the arguments so far, it is clear that the former type forms more preferable sensory associative patterns than the latter. In the former, the modifier is in a strong associative relation with the visual sense of reddishness, whereas the latter evokes the sensory association from the visual sense to the tactile sense, which is anomalous in our sensory experiences. The structure of sensory experiences plays a crucial role in synaesthesia.

The same principle seems to be operating in other senses such as the gustatory and olfactory senses. In daily life, we are likely to use the olfactory sense instead of the gustatory sense as a means to make a prediction of the property of an object, since there is more risk in using the gustatory sense than the olfactory sense. When the gustatory sense is used, the olfactory sense has normally already been used. The sensory association from the gustatory to the olfactory sense should therefore be stronger than the other way around. As well as the case of the tactile and visual senses, the adaptive behavior with the use of the gustatory and olfactory senses forms the structure of sensory experiences.

Expressions such as *sweet taste* are cross-linguistically acceptable (e.g., *tian2de wei4dao4* [Chinese: 1], *zoete smaak* [Dutch: 1], *γλυκειά γεύση* [Greek: 1], *amai aji* [Japanese: 1], *dan mas* [Korean: 1]). On the other hand, the acceptability of expressions such as *smelly taste* is likely to be low (e.g., *chou4de wei4dao4* [Chinese: 1], *stinkende smaak* [Dutch: -1], *δύσοσμη γεύση* [Greek: -1], *kusai aji* [Japanese: 0], *gulihan mas* [Korean: 0]). The acceptability of the former is higher than the latter, because the gustatory experience evoked by the modifier is in a strong associative relation with an olfactory experience. The latter type evokes the sensory association from the olfactory sense to the gustatory sense which is anomalous in the structure of sensory experiences.

Williams' generalization suggests that mapping from the olfactory sense onto the gustatory sense should not take place. The anomaly of *smelly taste* seems to follow his generalization. We suggest, however, that his suggestion is not accurate enough, since there is indeed a case where such a mapping is possible (e.g., *fen1fang1de wei4dao4* [Chinese: 1], *zoete smaak* [Dutch: 1], *fragrant taste* [English: 0], *ευωδιαστή γεύση* [Greek: 0], *kaguwasii aji* [Japanese: 0], *hyanggiloun mas* [Korean: 0]). Williams' generalization does not explain the different acceptability between *smelly taste* and *fragrant taste*.

In our accounts, frequencies of sensory co-occurrence determine the acceptability of synaesthesia. Use of the gustatory sense depends on the type of smell. For example, when an object has a good smell, one is more likely to eat it than when the object has a bad smell. Eating fragrant food forms the strength of the sensory association between the senses evoked by *fragrant* and *taste*. The acceptability of *fragrant taste* and the equivalent expressions across the languages is motivated by the strength of sensory association. The anomaly of *smelly taste* and its equivalent expressions stems from the lack of sensory association. The association between *smelly* and *taste* is weak, since eating smelly food is usually not welcome by people.

The principle of the structure of sensory experiences seems more accurate than Williams' generalization. Adaptive behavior contributes to the formation of the structure of sensory experiences, and the structure of sensory experiences determines the semantic structure of synaesthesia. Physiological properties of sense modalities and adaptive behavior are not limited to speakers of a particular language, and hence a similar pattern of possible synaesthetic mappings is found across languages. Below is the semantic constraint of synaesthesia we propose as a constraint of sensory experiences.

Constraint of Sensory Co-occurrence:

In a synaesthetic expression consisting of a modifier and a modified element, the modifier should cause a sensory reaction which is in a strong associative relation with the sense caused by the modified element.

4. Constraint of Emotional Experiences

Synaesthesia also includes examples such as *sweet voice*, where mapping from the gustatory sense onto the auditory sense takes place. Equivalent expressions to *sweet voice* are also likely to be found acceptable across the languages in our sample (e.g., *tian2de sheng1yin1* [Chinese: 1], *zoete stem* [Dutch: 0], *γλυκειά φωνή* [Greek: 1], *amai koe* [Japanese: 1], *dan mogsoli* [Korean: -1]). According to Williams' generalization, such a mapping is predicted. In our accounts provided so far, however, examples such as *sweet voice* would be predicted as unacceptable against the fact, since sweetness and a voice are not likely to co-occur in daily life. This section introduces another type of constraint and discusses the advantages of our arguments in comparison to Williams' work.

In emotional processing, the amygdala (a subcortical region in the temporal lobe) in the limbic system plays the crucial role, and the reactions of the amygdala are caused by sensory stimuli (Gazzaniga et al. 1998). We argue that comprehension of *sweet voice* is made possible not by the constraint of sensory experiences, but by the constraint of emotional experiences. Sweetness in the gustatory sense would cause an emotion of pleasure. Likewise, a certain type of voice causes a similar type of emotion. Comprehension of *sweet voice* is achieved by the synthesis of the similar emotional value of pleasure evoked by the lexical elements. (3) is a schematic description of the process.

(3) *Situation*:
Occurrence of the linguistic sounds of *sweet voice*
Reactions in the hearer:
- Reaction of the Auditory Sense [the linguistic sounds of *sweet*]
- Reaction of the Sensory Association [the association between the auditory sense of the linguistic sounds of *sweet* and the gustatory sense of sweetness]
- Reaction of the Gustatory Sense [sweetness]
- Reaction of the Sensory Association (the association between the gustatory sense of sweetness and the emotion of pleasure)
- Reaction of Emotion [emotion of pleasure]
- Priming in the Sensory Association [the association between the emotion of pleasure and the senses that are in strong associative relation with such an emotion]
- Reaction of the Auditory Sense [the linguistic sounds of *voice*]
- Reaction of the Sensory Association [the association between the auditory sense of the linguistic sounds of *voice* and the auditory sense of some kind of voice]
- Reaction of the Auditory Sense [a voice that causes an emotion of pleasure]

We proposed the constraint of sensory experiences in the preceding section. Synaesthesia involves two different types of constraints. The constraint of emotional experiences can be formalized as follows:

Constraint of Emotional Similarity:
In a synaesthetic expression consisting of a modifier and a modified element, the modifier should evoke an emotion which matches with the emotion evoked by the modified element.

There are some advantages to propose such a constraint. Firstly, in comparison with Williams' generalization, the Emotional Similarity constraint explains cases which cannot be captured by his generalization. For example, *bitter pain* and *sweet touch* conform to mapping from the gustatory onto the tactile sense, but such a fact cannot be captured by Williams' generalization. In our accounts, the acceptability of *bitter pain* is explained by the synthesis of the emotion of agony evoked by the lexical elements of this expression. Likewise, the acceptability of *sweet touch* is explained by the synthesis of the emotional value of pleasure evoked by *sweet* and *touch*.

Secondly, the constraint explains cases which are not assumed by Williams' generalization. For example, mapping from the olfactory onto the auditory sense is not assumed by Williams' generalization, but there are in fact instances such as *fragrant music*, which conforms to such a mapping. In our account, comprehension of *fragrant music* is made possible by the synthesis of pleasant feeling evoked by the lexical elements. In daily life, fragrance and sound are not likely to co-occur. It is not the Sensory Co-occurrence constraint, but the Emotional Similarity constraint that makes such a mapping possible.

Thirdly, the constraint is consistent with anatomical structures of the brain. The amygdala is interconnected with various structures in the limbic system, and one of these includes the hypothalamus (an output system for the amygdala) (Rolls 1999). An interesting fact about the hypothalamus is that hypothalamic neurons in the lateral hypothalamus are responsive to the sight, smell, and taste of food (Rolls 1999: 19). According to the results of our survey, a number of examples that violate Williams' generalization relate to the visual, olfactory, and gustatory senses (see Figure 2). In our accounts, such linguistic facts and the empirical findings of the brain can also be captured in relation to emotion.

Moreover, in relation to the above discussion, the constraint also captures subtle aspects of synaesthesia. An emotional value caused by sensory stimulus should differ among individuals or in cultures. Figure 2 (d), for example, shows that Japanese shows a peculiarity to other languages in terms of the mapping from the auditory sense to the gustatory sense. In Japanese, for example, *urusai aji* is judged acceptable if given a context, but its equivalent expressions in other languages are likely to be judged anomalous (e.g., *chao3za2de wei4dao4* [Chinese: -1], *luide smaak* [Dutch: -1], *loud taste* [English: -1], *δυνατή γεύση* [Greek: 1], *solanseuleoun mas* [Korean: -1]). In our accounts, such differences are explained by the availability of emotional matching.

Last but not least, the constraint makes it clear what kind of principle causes a mapping to be possible. In section 3, we mentioned that Williams' generalization does not tell us the cause of the different acceptability between *smelly taste* and *fragrant taste*. A similar ambiguity can be found in instances such as *sweet voice*. As mentioned, his generalization predicts *sweet voice* to be possible, but it is again not clear why it is so. In our accounts, it is the Emotional Similarity constraint that makes such a mapping possible.

5. Conclusion

This paper introduced a cognitive model called the PP model and discussed constraints of synaesthesia across a variety of languages. Two types of constraints were proposed. Regarding the constraint of sensory experiences, we discussed the importance of what we call the structure of sensory experiences. Regarding the constraint of emotional experiences, we mentioned the role played by emotion. We showed that Williams' generalization suffers from several problems. By using the PP model, we presented a systematic analysis of synaesthesia which is compatible with the structure of the brain.

References

Gazzaniga, M. S., R. B. Ivry, & G. R. Mangun. 1998. *Cognitive neuroscience: the biology of the mind.* New York: W.W. Norton.

Rolls, Edmund T. 1999. *The brain and emotion.* Oxford: Oxford University Press.

Shibuya, Yoshikata, Hajime Nozawa, & Toshiyuki Kanamaru. in press. Understanding synaesthetic expressions. In Peter Holz & Martina Plümacher (eds.) *How can language cope with odors and colors?* Amsterdam: John Benjamins.

Williams, Joseph M. 1976. Synaesthetic adjectives: a possible law of semantic change. *Language* 52: 461-478.

Yamanashi, Masa-aki. 1988. *Hiyu to rikai* (Metaphor and understanding). University of Tokyo Press.

Yoshikata Shibuya
Department of Linguistics
University of Manchester,
Oxford Road,
Manchester, M13 9PL, England

yoshikata.shibuya@stud.man.ac.uk

Hajime Nozawa
Graduate School of Human and
Environmental Studies
Kyoto University,
Nihonmatsu-cho, Yoshida, Sakyo-ku,
Kyoto 606-8501, Japan

hajime_nozawa_jp@yahoo.co.jp

Reexamining Condition C Effects[*]

KAYONO SHIOBARA
University of British Columbia

0. Introduction

This paper reexamines Binding Condition C effects and shows that Binding Condition C is not a reliable diagnostic for the distribution of nominal phrases. The main claim of this paper is that Condition C effects should be captured in terms of a condition on anaphoric dependence ("General Patterns of Anaphoric Dependence" in Williams 1997), not by a structural condition on coreference ("Binding Theory Condition C" in Chomsky 1981).

1. Condition C effects as anaphoric dependence effects

When a sentence contains an R-expression that serves as the antecedent of another NP, the distribution of the NPs is constrained in a certain way. For example, the sentences in (1) are acceptable when the R-expression *his term paper* precedes the pronoun *it* as in (1a)-(1b), but when the pronoun follows the R-expression, the sentence is acceptable only if the pronoun is within a subordinate clause as in (1c). If the pronoun is within a matrix clause, the sentence is unacceptable as in (1d).[1]

(1) a. Anyone [who has written **his term paper**] can turn *it* in to me now.
    b. Anyone can turn **his term paper** in to me now [who has written *it*].
    c. Anyone [who has written *it*] can turn **his term paper** in to me now.
    d. *Anyone can turn *it* in to me now [who has written **his TERM PAPER**].
(Williams 1997:587)

Binding Theory Condition C (BT-C), defined as in (2), attributes the unacceptability of the sentence in (1d) to the R-expression *his term paper* which is illegitimately bound by (i.e. c-commanded by and coindexed with) the pronoun *it*.

[*]I thank the members of UBC pronoun reading group, particularly Rose-Marie Déchaine and Martina Wiltschko, for valuable comments on earlier versions of this paper. Thanks also go to Lynn Nichols and Benjamin Shaer for insightful comments and questions at and since the annual meeting of BLS. The research in this paper is in part supported by SSHRC 410-2002-1078 to Martina Wiltschko (principle investigator), which is gratefully acknowledged.

[1]Throughout the paper, the XP which serves as an antecedent of a proform is **bolded**, the proform is in *italics*, and the word(s) with main sentence stress is(/are) in UPPER CASE LETTERS.

(2) Binding Theory Condition C (BT-C)
An R-expression is free. (Chomsky 1981:188)

On the other hand, Williams (1997) takes the pattern in (1) as an instance of the General Patterns of Anaphoric Dependence (GPAD) which is illustrated as follows:

(3) General Patterns of Anaphoric Dependence (GPAD, based on Williams 1997:588)

forward dependence	A.	[...antecedent...]subord	[...pro...]matrix
	B.	[...antecedent...]matrix	[...pro...]subord
backward dependence	C.	[...pro...]subord	[...antecedent...]matrix
	D.	*[...pro...]matrix	[...antecedent...]subord

The (3A)-(3D) patterns correspond to the (a)-(d) examples in (1). The GPAD says that anaphoric dependence can be forward as in (3A-B), or backward and down as in (3C). The dependence of a pronoun on its antecedent NP in (1) is not the only paradigm governed by the GPAD and examples that involve a null VP and its antecedent are also subject to the GPAD:

(4) a. Anyone [who wants to **see the doctor**] can 0_{VP}.
b. Anyone can **see the doctor** [who wants to 0_{VP}].
c. Anyone [who wants to 0_{VP}] can **see the doctor**.
d. *Anyone can 0_{VP} [who wants to **see the doctor**]. (cf. Williams 1997:591)

The GPAD is designed to rule in an acceptable sentence such as (5), where the putative antecedent does not have main sentence stress and the antecedent itself becomes anaphoric:

(5) [Context: I assume you recall that this course requires **a term paper**.]
Anyone can turn *it* in to me now [who has WRITTEN his term paper].
(Williams 1997:588)

In this case, the pronoun *it* is dependent not on the following NP *his term paper*, but on the preceding NP *a term paper* in the context, thereby canceling the illegitimate backward-up pattern in (3D) (cf. compare with (1d)). Note further that c-command is not the relation governing the paradigm in (1), and the BT-C does not predict the unacceptability of (6):

(6) *Anyone can try [to hand *it* in to me] [who has written **HIS TERM PAPER**].
(Williams 1997:588)

In (6), the pronoun *it* is embedded more deeply than in (1d) and the R-expression cannot be c-commanded by the pronoun. Still, the sentence is unacceptable. The GPAD correctly excludes the sentence in (6) because this is the illegitimate backward-up pattern.

In the next section, I will show that Condition C effects in general are better captured in terms of the GPAD than the BT-C.

2. Deriving Condition C effects
2.1. Forward dependence: wh-movement
An apparent problem for the GPAD is found in wh-movement examples:

(7) a. Which claim [that offended **John**] was *he* willing to talk about _?
 b. *Which claim [that **John** was wrong] was *he* willing to talk about _?

The contrast between a wh-phrase with an adjunct relative clause and a wh-phrase with a complement clause with respect to the interpretational possibilities of the name *John* is observed in e.g. Riemsdijk and Williams (1981), Lebeaux (1988,1991): if the name is contained in an adjunct as in (7a), the name can refer to the same entity as the following pronoun *he*, but this interpretation is not available if the name is contained in a complement as in (7b). That is to say, "Condition C effects are abrogated when the fronted name is contained in an adjunct" (Lebeaux 1991:212).

Before looking at the proposed analysis of this contrast, let me talk about the empirical issues regarding the contrast. First, the contrast in (7) is not as clear as it was originally claimed, which is suggested by acceptability judgments by three native speakers I consulted:

(8) Acceptability of the sentences in (7)

	Speaker A	B	C
(7) a.	√	√	??
b.	?	?	?*

Furthermore, Speaker A notes that the sentence in (7b) is as good as (7a) if there is a preceding context that "several people made a claim that John was wrong". Also, the contrast in (7) is questioned in the literature and there are more than a few acceptable wh-movement sentences with a complement clause, two of which are presented in (9).

(9) a. Which piece of evidence [that **John** was guilty] did *he* successfully refute? (Lasnik 1999:209)
 b. Whose claim [that **the Senator** has violated the campaign finance regulations] did *he* dismiss as politically motivated? (Kuno 1997)

With this in mind, let us look at the analyses.

2.1.1. GPAD
The GPAD predicts that all the wh-movement sentences in (7) and (9) are acceptable because they all show the legitimate forward-down pattern in (3A).

(7') a. ... [...**John**...] ...*he*... : GPAD (3A)
 b. *... [...**John**...] ...*he*... : GPAD (3A)
(9') a. ... [...**John**...] ...*he*... : GPAD (3A)
 b. ... [...**the Senator**...] ...*he*... : GPAD (3A)

Since the GPAD is not sensitive to the distinction between adjunct and complement clauses, it does not predict the contrast in (7). Given that (7b) is not as bad as claimed and becomes acceptable in an appropriate context, it does not constitute a counterexample to the GPAD.

2.1.2. Binding Theory Condition C: Lebeaux's explanation

The acceptability of the sentence in (7a) is referred to as the "anti-reconstruction effect" in the sense that the wh-phrase that contains the name *John* does not seem to reconstruct into the pre-wh-movement position (indicated by the underscore in the examples repeated below) where Binding Condition C (BT-C) would be violated.

(7) a. Which claim [that offended **John**] was *he* willing to talk about _?
b. *Which claim [that **John** was wrong] was *he* willing to talk about _?

Lebeaux (1988,1991) explains the contrast in terms of the difference in how adjuncts and complements enter the derivation. He argues that the Projection Principle (Chomsky 1981:29) requires complements, but not adjuncts, to be present at D-structure, and adjuncts could be attached at any stage in the derivation. Given this argument, the sentence in (7a) can have a derivation such as (10a), whereas (7b) has the derivation in (10b):

(10) a. Derivation of (7a)

he was willing to talk about which claim that offended **John** ⇒ wh-movement

[which claim] was *he* willing to talk about _ that offended **John** ⇒ late-merge of adjunct

[which claim [that offended **John**]] was *he* willing to talk about _?
: BT-C satisfied

(10) b. Derivation of (7b)

he* was willing to talk about which claim that **John was wrong
: BT-C violated
⇒ wh-movement

[which claim that **John** was wrong] was *he* willing to talk about _?

In (10a), wh-movement occurs before the adjunct clause is adjoined (i.e. "merged" in Minimalist terms, cf. Chomsky 1995) into the wh-phrase, and the adjunct clause is adjoined to the already fronted wh-phrase via "late-merge", thereby satisfying the BT-C at every stage of the derivation. In contrast, the complement clause in (10b) must be merged into the wh-phrase at D-structure, resulting in violation of the BT-C, which Lebeaux assumes to apply throughout the derivation.

There are empirical and theoretical problems with Lebeaux's proposal regarding the contrast in (7). First, BT-C under his mechanism excludes the cases of wh-phrase with a complement altogether, and hence it wrongly excludes acceptable sentences in (9) and also (7b) when it is acceptable with an appropriate context. Theoretically, Lebeaux's rationale for exceptional behavior of adjuncts is not available in the standard Minimalist framework (Chomsky 1995) where D-structure and the Projection Principle are abandoned. Furthermore, the late-merge operation does not satisfy the Extension Condition on structure-building in that an adjunct is embedded within another syntactic object.

2.2. Backward dependence: extraposition from NP

Fox and Nissenbaum (1999) argue that there is a contrast with respect to Condition C effects between extraposition of an adjunct and extraposition of a complement:

(11) a. I gave *him* an argument yesterday [that supports **John**'s theory].
b. ??/*I gave *him* an argument yesterday [that this sentence supports **John**'s theory]. (Fox and Nissenbaum 1999:138)

Their generalization is that "extraposition of adjuncts bleeds Binding Condition C" (p.139). As is the case with wh-movement sentences, however, the contrast between adjunct and complement cases is subtle if any, and three native speakers I consulted find the sentence in (11b) as bad as the sentence in (11a). Furthermore, both (11a) and (11b) become acceptable when the putative antecedent *John* is destressed and used anaphorically in an appropriate context:

(12) [Context: **John** has been facing lots of counterexamples to his theory and is quite disappointed at that. But finally…]
a. ?I gave *him* an argument yesterday [that SUPPORTS John's theory].
b. ?I gave *him* an argument yesterday [that THIS SENTENCE supports John's theory].
b'. ?I gave *him* an argument yesterday [that this sentence SUPPORTS **John**'s theory].

With these data in mind, let us turn to the analysis.

2.2.1. GPAD

The GPAD predicts that the extraposition from NP sentences in (11) are unacceptable because they represent the illegitimate backward-up pattern in (3D).

(11') a. …*him*… […**John**'s…] : GPAD *(3D)
b. ??/* …*him*… […**John**'s…] : GPAD *(3D)

Given that the acceptability of (11a) is dubious, it does not constitute a counterexample to the GPAD.

The improvement of the sentences in an appropriate context, illustrated in (12), is exactly what the GPAD predicts.

(12') a. ?**John**… …*him*… […John's…] : GPAD *(3D) cancelled
b. ?**John**… …*him*… […John's…] : GPAD *(3D) cancelled
b'. ?**John**… …*him*… […John's…] : GPAD *(3D) cancelled

In (12), the putative antecedent NP *John* is used anaphorically and hence destressed. Therefore, the pronoun *him* is dependent on the NP *John* in the preceding context, not on the following NP *John* within the same sentence, the illegitimate backward-up pattern being cancelled.

2.2.2. Binding Theory Condition C: Fox and Nissenbaum's explanation

Fox and Nissenbaum (1999) propose different derivations for extraposition of adjuncts and extraposition of complements. In addition to Lebeaux's proposal that adjuncts, but not complements, can late-merge, they assume that a host NP of an extraposed adjunct undergoes rightward Quantifier Raising (QR) to the vP-adjoined position, and spell-out can apply more than once. The derivations of the extraposition from NP sentences in (11) are illustrated in (13):

(13) a. Derivations of (11a)

I gave *him* an argument yesterday that supports **John**'s theory ⇒ Spell-out, and QR of the host NP

I [$_{vP}$ [$_{vP}$ gave *him* [an argument] yesterday] [~~an argument~~]] that supports **John**'s theory ⇒ late-merge of adjunct

I [$_{vP}$ [$_{vP}$ gave *him* [an argument] yesterday] [~~an argument~~ [that supports **John**'s Theory]]]. ⇒ Spell-out : BT-C satisfied

b. Derivation of (11b)

I gave *him* [an argument [that this sentence supports **John**'s theory]] yesterday. ⇒ extraposition from NP

*I gave *him* [an argument _] yesterday [that this sentence supports **John**'s theory]. : BT-C violated

First, let us look at the derivation of extraposition of an adjunct in (13a). With Lebeaux's proposal, the adjunct relative clause can start as an independent syntactic object. After the root sentence gets spelled out, the host NP *an argument* undergoes QR to the vP-adjoined position, and the adjunct is late-merged into the QR'ed NP. Since the host NP is already spelled out, the QR'ed NP is not pronounced. There is no BT-C violation because the name *John* is in a higher position than the object NP *him*. In the case of extraposition of a complement in (13b), neither late-merge nor QR of the host NP is available and the complement clause undergoes rightward extraposition movement in a traditional fashion, presumably to the position c-commanded by the object NP, where the BT-C is violated.[2]

Besides the empirical issues regarding the contrast in (11), i.e. the contrast is not as clear as Fox and Nissenbaum claim, and the sentence in (11b) significantly improves in

[2]Fox and Nissenbaum (1999) are not clear about the exact landing site of extraposed complements.

an appropriate context as is shown in (12b), there are conceptual and theoretical problems with Fox and Nissenbaum's explanation. First, the QR operation of a host NP in extraposition of adjuncts is not independently motivated.[3] Moreover, nothing guarantees that QR should be "rightward" (as is pointed out in Chomsky 2001:18-19). Their explanation also carries over Lebeaux's theoretical problem that late-merge is an extraordinary structure-building operation in that it does not satisfy the Extension Condition.

2.3. Another argument against Binding Theory Condition C: *as*-parentheticals

It is observed that Condition C effects in sentences with *as*-parentheticals are far from clear:

(14) a. ?*She* owns, as **Sue** told us, a dozen or so unicycles.
b. **He* told *his* mom that, as **Julio** had promised, the dishes were done.
(Potts 2002:666fn)

The variation in acceptability is independent of the issue of adjunct vs complement distinction, because there is an agreement that *as*-parentheticals are not complements to any head no matter how they are represented in syntax. The dependence pattern in the sentences in (14) is backward, and whether it is up or down depends on how *as*-parentheticals are analyzed syntactically. The GPAD predicts that the sentence in (14b) improves if the pronouns (*he* and *his*) take an antecedent from the preceding context, thereby achieving forward dependency. This prediction is borne out to a certain extent:

(15) [Context: **Julio** does not do the dishes and is always scolded by his mom. Yesterday his mom had to stay at work until late and made **Julio** promise her to do the dishes by the time she comes back home. When she got home…]
?/**He* told *his* mom that, as Julio had promised, the dishes were done.

Among two native speakers who did not accept the sentence in (14b), one of them accepted the sentence in the context in (15), but the other did not. To interpret this individual variation, more examples from *as*-parentheticals should be checked with native speakers, which I leave for future research.

3. Consequences of the proposal

3.1. Eliminating "late-merge" from the computational component

By having the Binding Theory Condition C (BT-C) being replaced with the General Patterns of Anaphoric Dependence (GPAD) in explaining Condition C effects, the computational component does not have to resort to the late-merge operation any more. This is a welcome consequence theoretically (cf. Chomsky 2001:19) because late-merge departs from the Minimalist way of structure-building (i.e. Merge) in that it does not satisfy the Extension Condition on structure-building. The elimination of late-merge was made possible only after reexamining the alleged contrast between adjunct and

[3] Fox and Nissenbaum (1999) show that presence and absence of QR in adjunct extraposition and complement extraposition, respectively, *can predict* their difference in scope-marking, but still fail to *motivate* QR in adjunct extraposition.

complement displacement cases by embedding them in appropriate contexts and abandoning the significance of those contrasts (e.g. wh-movement in (7) and extraposition from NP in (11)).[4]

3.2. Loose ends

Although the GPAD is designed to apply generally to any anaphoric dependence, it is too weak in that it wrongly allows unacceptable dependence cases. For example, nothing should be wrong with the unacceptable sentences in (16) as far as the GPAD is concerned because they establish legitimate forward-down (3A) and backward-down (3C) patterns, respectively:

(16) a. *[Which claim [that offended **everyone**]] was *he* willing to talk about [~~which claim [that offended everyone~~]]?
a'. *[...**everyone**...] ...*he*... : GPAD (3A)
b. [Which claim [that offended *him*]] was ***everybody**/??**every man/each man** willing to talk about [~~which claim [that offended *him*~~]]?
b'. */??/√[...*him*...] ...**QP**... : GPAD (3C)

I argue here that the sentence in (16a) is excluded by the c-command condition on bound variables, independently of the GPAD: the pronoun *he* is not bound by the quantifier phrase (QP) *everyone* at any point of the derivation, and hence cannot have a bound variable reading. Now that we abandon late-merge of adjuncts, nothing prevents the existence of the adjunct clause in base position. However, the QP within the adjunct in the base position does not c-command the pronoun either. On the other hand, the acceptability of the sentence in (16b) varies depending on the type of the antecedent QP. The acceptability of the sentence with the QP *each man* is compatible with the GPAD, and I argue that the unacceptability of the sentences with the QPs *everybody* and *every man* is due to a condition other than the GPAD (the nature of which is still to be examined).

Related to the empirical coverage of GPAD, it is worth noting here the theoretical issues and problems regarding the nature of GPAD. The first problem is that the GPAD is too weak in the sense that we have just seen and needs to be supplemented by other conditions on the distribution of NPs, for example, the c-command condition on bound variable pronouns.

Another major issue is why the GPAD is stated the way it is. Unless we answer this question, the GPAD is merely a description of linguistic phenomena, not an explanation. The GPAD consists of two cross-cutting notions: forward vs backward and up vs down. My speculation is that these notions are motivated by the properties of human cognitive processes, in particular, sentence processing. For example, it is easier to process the dependence relation if the antecedent with an independent semantic content precedes its dependent than the opposite pattern (cf. Hawkins 2002) and the opposite pattern is possible to the extent that the structural relation between an antecedent and its dependent

[4] When it comes to PP complements and adjuncts, things are even less clear. (See e.g. Reinhart 1983 and Speas 1990 for Condition C effects observed in fronted PPs.) Shaer (p.c.) notes that in German, fronted PPs, whether they are adjuncts or complements, usually show Condition C effects (i.e. no anti-reconstruction effects).

is compatible with a certain cognitive (processing) domain (cf. Reinhart 1983:ch10, Deane 1992). This issue remains to be further explored.

4. Summary

The General Patterns of Anaphoric Dependence (GPAD) proposed in Williams (1997) captures Condition C effects better than the Binding Condition C (BT-C), in that the former (i) rules in acceptable examples which the latter would wrongly rule out, and (ii) predicts variations in acceptability judgments. A welcome consequence of replacing the BT-C with the GPAD is that the theoretically problematic operation, late-merge, can be eliminated. The GPAD, however, lacks explanatory power at the moment.

References

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Berlin: Mouton de Gruyter.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Beyond explanatory adequacy. Ms., MIT.
Deane, Paul D. 1992. *Grammar in Mind and Brain: Explorations in Cognitive Syntax*. Berlin: Mouton de Gruyter.
Fox, Danny and Jon Nissenbaum. 1999. Extraposition and scope: a case for overt QR. *WCCFL* 18: 132-144.
Hawkins, John A. 2002. Symmetries and asymmetries across grammars: some predictions from parsing. *Theoretical Linguistics* 28: 95-150.
Kuno, Susumu. 1997. Binding theory in the minimalist program. Ms., Harvard University.
Lasnik, Howard. 1999. Chains of arguments. In S. D. Epstein & N. Hornstein (eds.) *Working Minimalism*, 189-215. Cambridge, MA: MIT Press.
Lebeaux, David. 1988. *Language acquisition and the form of the grammar*. Ph.D. dissertation, University of Massachusetts, Amherst.
Lebeaux, David. 1991. Relative clauses, licensing, and the nature of the derivation. *Syntax and Semantics* 25: 209-239.
Potts, Chris. 2002. The syntax and semantics of *as*-parentheticals. *Natural Language and LinguisticTheory* 20: 623-689.
Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. London: Croom Helm.
Riemsdijk, Henk van and Edwin Williams. 1981. NP-structure. *Linguistic Review* 1: 171-217.
Speas, Margaret. 1990. *Phrase Structure in Natural Language*. Amsterdam: Kluwer Academic Publisher.
Williams, Edwin. 1997. Blocking anaphora. *Linguistic Inquiry* 28: 577-628.

Department of Linguistics
E270-1866 Main Mall
University of British Columbia
Vancouver, B.C. V6T1Z1
Canada

kayono@interchange.ubc.ca

Analogical Modeling: Exemplars, Rules, and Quantum Computing

ROYAL SKOUSEN
Brigham Young University

1. Properties of analogical modeling

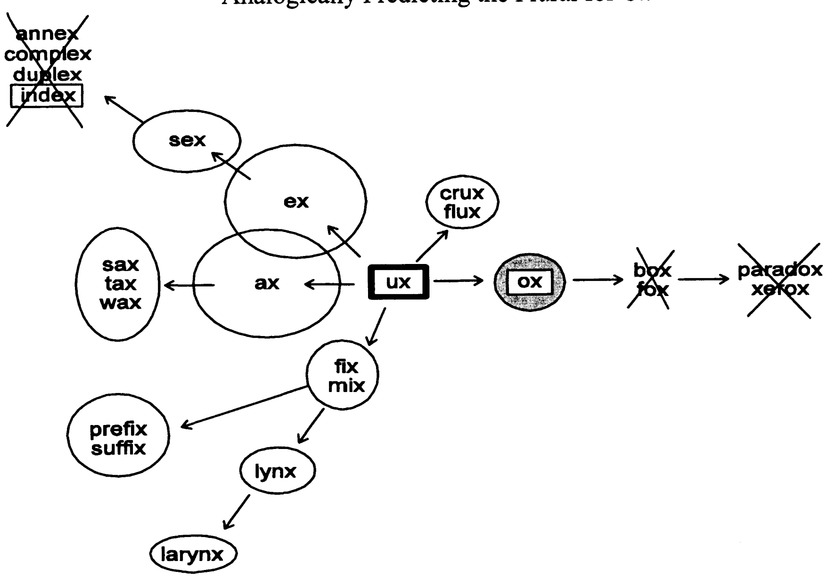
I begin this paper by outlining the main properties of analogical modeling (AM):

- AM is an exemplar or instance-based system of prediction; it is not a rule-based system nor a neural network.
- AM is a procedural system, not a declarative one; one cannot directly find its predictions within the system itself. Instead, predictions are always made in terms of a given context for which a prediction (or outcome) is sought.
- AM is not a nearest neighbor approach; it does include nearest neighbors in its predictions, but it also regularly uses non-neighbors to make predictions. However, such non-neighbors can be used only under a well-defined, explicit condition of homogeneity.
- No training stage occurs in AM, except in the trivial sense that one must collect exemplars in order to make predictions. There is no setting of parameters nor any prior determination of variable significance. The significance of any combination of variables is always determined in terms of the given context for which we seek a predicted outcome. Each prediction is done "on the fly".
- AM has a simple measure of uncertainty that is defined by the number of outcome disagreements between pairs of exemplars.
- AM uses a simple decision principle to determine homogeneity: namely, minimize the number of disagreements. No increase in uncertainty is permitted, which means that no analogical analysis will allow any loss of information.
- This definition of homogeneity results in two types of homogeneous spaces in an analogical analysis: either (1) the space is *deterministic* in behavior, which means that there will be no disagreements at all within that space and therefore no a priori limitation on where occurrences are found within the deterministic space; or (2) the space is *non-deterministic* in behavior, which means that even though disagreement occurs within that space, its occurrences are all restricted to a single subspace within the non-deterministic space.
- The resulting probability of using a particular exemplar depends upon three

factors: (1) *proximity:* the closer the exemplar to the given context, the greater its chances of being selected as the analogical model; (2) *gang effect:* when a group of exemplars in the same space behave alike, the chances of one of those exemplars being selected is multiplied; (3) *heterogeneity:* the chances of an exemplar being used is zero whenever there is any intervening exemplar closer to the given context that behaves differently.

We can see these three properties in the following schematic. Here we have a class of possible exemplars for analogically predicting the English plural of the nonce noun *ux*. For purposes of illustrating the properties of analogical modeling, the dataset for predicting the given context (*ux*) is arbitrarily restricted to nouns ending in the spelling *x*. There are 23 exemplars in the dataset, of which most take the regular plural, *-es*, but two have exceptional plurals: (1) the nearest-neighbor *ox*, which normally takes the irregular ending *-en* (thus, *ox*/*oxen*) but infrequently is regularized (*ox*/*oxes*), and (2) the distant *index*, which has two plurals, each about equally probable: the regular *indexes* and the Latinate *indices*.

Analogically Predicting the Plural for *Ux*



In this schematic, the given context (*ux*) is in a heavy square box. The given context is surrounded by its nearest neighbors, *ax, ex, crux, flux,* and *ox*. The size of the circles surrounding each group of neighbors approximately shows the chances of selecting each of these exemplars as the analogical model for predicting the plural for *ux*. The circles for *ax* and *ex* are the largest because they are supported by various gangs of less-near neighbors that behave like *ax* and *ex* (namely *sax, tax, wax, sex, fix, mix, lynx, larynx, prefix,* and *suffix*). The further we go out from the given context, though, the smaller the circles become, thus showing a decrease in choosing less-near neighbors because of their relative lack of proximity to *ux*. Although *crux* and *flux* are nearest neighbors, their strength is less because they have no support from less-near neighbors.

In contrast, the behavior of *ox* leads to heterogeneity in the contextual space. There is some chance that *ox,* a nearest neighbor, will serve as the analogical model for *ux*, which would predict the plural *uxen.* But there are no other exemplars that will support *ox* in predicting *uxen.* Even regularly behaving exemplars like *box, fox, paradox,* and *xerox* will have no influence on predicting the plural for *ux* since the differently behaving *ox* is closer to *ux* and stands between the given context and those less-near exemplars. For this reason, these four outliers, even though they are regular in behavior, occur in a heterogeneous subspace of the contextual space and are therefore totally excluded from the analogical set. Of course, there are enough regular exemplars still in the analogical set to allow the regular behavior to be overwhelmingly predicted. But because of the proximity of *ox* to the given *ux*, there remains a small, but noticeable, possibility that the predicted plural will be *uxen.*

Finally, we note that there is a group of distant exemplars that are also excluded from the analogical set for *ux,* namely *annex, complex, duplex,* and *index*. The nondeterministic behavior of *index* (with its varying plural *indexes* versus *indices*) means that the subspace containing all four of these words is heterogeneous. Again the reason is that there are intervening exemplars closer to the given *ux* that behave differently (namely, *ex* and *sex* behave differently from *index*, at least for exemplars taking *indices* as the plural). And because *index* is therefore excluded, all the other exemplars in the same subspace (namely *annex, complex,* and *duplex*) are also eliminated from the analogical set, even though their own behavior is as regular as the intervening exemplars *ex* and *sex*. Thus we have a second example of some regularly behaving exemplars being excluded because of heterogeneity.

2. The basic literature

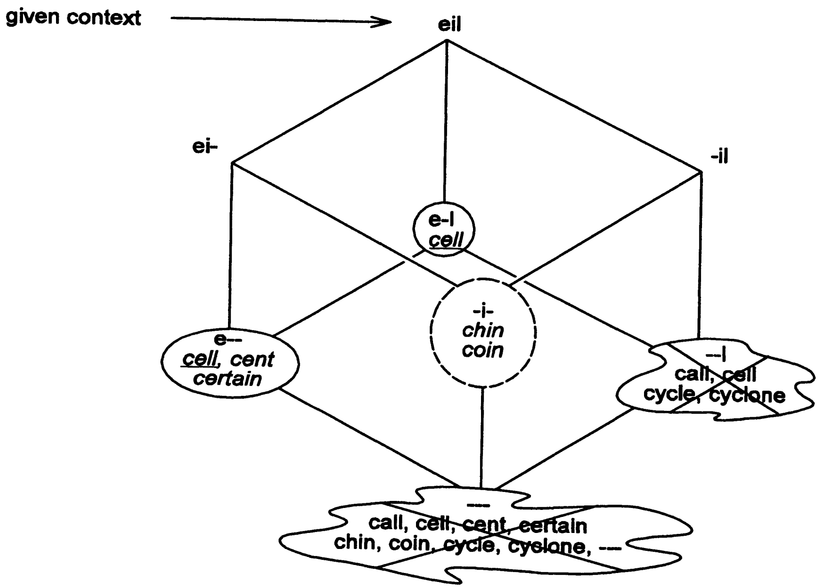
Analogical modeling is treated in three fundamental books: (1) *Analogy and Structure* (Skousen 1992), (2) *Analogical Modeling of Language* (Skousen 1989), and (3) *Analogical Modeling: An exemplar-based approach to language* (Skousen, Lonsdale, and Parkinson 2002).

Written in the early 1980s and finished in 1984, *Analogy and Structure* describes the fundamental properties of both rules and analogy and provides the mathematical basis for the theory of analogical modeling. This book was actually published later than *Analogical Modeling of Language,* which was written in the mid 1980s to provide a more general outline of the theory of analogical modeling and, in particular, its application to language. The third book, *Analogical Modeling: An exemplar-based approach to language*, describes current developments in the theory. Besides providing a tutorial on AM and how to run the computer program, it describes the psycholinguistic evidence for AM and applies the theory to a number of specific language problems. The theory is also compared to nearest neighbor approaches. Quantum computing is proposed as the natural way to handle the computational exponential explosion that is inherent in analogical modeling.

3. The terminology of analogical modeling

In this section I provide a sample analogical set in order to explain the terminology of AM. In the following example, I use AM to predict the pronunciation of the initial *c* letter in *ceiling*. Our DATASET will be a highly restricted one consisting of 34 words with an initial *c* letter. (The words in this simple dataset can be found in Skousen, Lonsdale, and Parkinson 2002:13-14.) In this dataset, there are three possible OUTCOMES: *k*, *s*, and *č* (based on words like *call*, *cent*, and *chin*). For this particular simplified example, we will try to predict the pronunciation of the initial *c* in *ceiling* by using the first three following letters (after the initial *c*) as VARIABLES. We make our predictions in terms of a GIVEN CONTEXT, which for *ceiling* will be $e\ i\ l$. The remaining letters $i\ n\ g$ will be ignored in this analysis. With three variables, there are eight possible combinations of variables that can be used to predict the outcome, namely $e\ i\ l$, $e\ i\ -$, $e - l$, $- i\ l$, $e - -$, $- i -$, $- - l$, and $- - -$. Whenever the dash ($-$) is used, it means that the actual particular value of that variable is being ignored, or in other words, that variable can be assigned anything. These combinations of variables are called SUPRACONTEXTS. In general, if there are n variables, there will be 2^n supracontexts.

The 2^n supracontexts form a partial ordering based on subcontextual relationships; that is, we connect more general supracontexts to more specific supracontexts, providing the more specific ones are subcontexts of the more general ones. And for each supracontext we list the dataset occurrences found in that supracontext (see the diagram on the next page). Three of the supracontexts, including the given context $e\ i\ l$, have no occurrences and are thus empty. Three of the supracontexts, all encircled, are HOMOGENEOUS in behavior. At least one of two conditions must be met for a supracontext to be homogeneous: either there is only one outcome type in the supracontext or all the occurrences are found in the same subcontextual subspace. In the first case, the supracontext is deterministic in behavior since only one outcome is found. Two of the homogeneous supracontexts are deterministic, namely $e - l$ and $e - -$. The occurrences in these two supracontexts take only the *s* outcome. I encircle these two deterministic supracontexts with a solid line.



The other homogeneous supracontext $-i-$ is non-deterministic in behavior. There are two different outcomes in this supracontext (namely $\check{c}$ for *chin* and k for *coin*). But these two occurrences are found only within this supracontext and not in any more specific supracontext higher up in the lattice. In the above diagram, this non-deterministic homogeneous supracontext is encircled with a dashed line. Basically, a non-deterministic homogeneous supracontext permits no other occurrences between itself and the (non-occurring) given context. A deterministic homogeneous supracontext, on the other hand, does allow such intervening occurrences. For instance, the more general supracontext $e--$ contains three occurrences, of which one, *cell,* is also found in $e-l$ and thus closer to the given context than the two other occurrences, *cent* and *certain*. But all three of these have the same outcome s, thus homogeneity is maintained. It should also be noted here that *cell* is found in two different homogeneous supracontexts, namely $e-l$ and $e--$. In a sense, we can think of *cell* as multiply occurring. This property of multiple existence has important consequences when we compare analogical modeling to quantum mechanics.

Finally, two of the supracontexts turn out to be HETEROGENEOUS, namely $--l$ and the most general supracontext $---$. Both are non-deterministic and their occurrences are not restricted to a single subcontext. For instance, $--l$ contains *call* (with outcome k) and *cell, cycle,* and *cyclone* (with outcome s), which makes this supracontext non-deterministic. But in addition, *cell* is found in $e-l$, which is closer to the given context and thus there is an intervening different behavior. In a similar way, the most general supracontext is also heterogeneous. In fact, it is easy to show that if a supracontext is heterogenous, a more general one containing that one as a subcontext will also be heterogeneous. We refer to this property as INCLUSIVE HETEROGENEITY. In other words, $---$ is inclusively heterogeneous because the more specific supracontext $--l$ is heterogeneous. In the above lattice, we enclose each heterogeneous supracontext with a jagged line and also place a large X through each one, thereby indicating that none of its occurrences can be accessed.

Occurrences that are far enough away from the given context are eliminated as possible exemplars. Closer occurrences can be found in homogeneous supracontexts, and those occurrences can be used as exemplars. Note, in particular, that the same occurrence can be inaccessible from a heterogeneous supracontext, but accessible from a homogeneous supracontext. Thus *cell* is accessible from both $e-l$ and $e--$ (which are homogeneous), but not from $--l$ or $---$ (which are heterogeneous).

We now use the resulting analogical set to predict the outcome. Instead of directly selecting an occurrence in a homogeneous supracontext, we rely on directional pointers that connect every pair of occurrences in a homogeneous supracontext. The basic RULE OF USAGE for predicting behavior is to randomly select one of these pointers and to observe which occurrence it is pointing to. This occurrence becomes the analogical model and the outcome associated with this exemplar is selected as the predicted outcome. We refer to this rule of usage as RANDOM SELECTION. By selecting pointers, the probability of choosing a particular homogeneous supracontext is the square of the frequency of that supracontext. Underlyingly, the frequency of the supracontext is a linear measure and is equal to the number of occurrences in the supracontext, but the number of pointers is equal to the frequency squared. Another possible rule of usage is SELECTION BY PLURALITY, which allows us to examine the occurrences predicted by random selection and choose the one that is most frequent. This rule of usage leads to the best possible decision whenever we are trying to maximize gain or minimize loss.

The analogical set can now be used to predict the pronunciation of the initial *c* in the word *ceiling*. For each homogeneous supracontext, we list the number of occurrences for each outcome, plus the number of pointers to each outcome. (See the chart on the next page.) If the underlying frequency (or number of occurrences) for a homogeneous supracontext is m, then there will be m^2 pointers, proportionally distributed. Empty supracontexts will have no occurrences or pointers and are trivially homogeneous. Every heterogeneous supracontext (marked by an x) may have occurrences, but none of their pointers will be accessible. For random selection, the probability of selecting the *s* outcome dominates, even though we only had 34 sample spellings in our restricted dataset. Using selection by plurality, the *s* outcome would be chosen every time.

supracontexts		*occurrences*			*pointers*		
		k	*s*	*č*	*k*	*s*	*č*
	e i l	-	-	-	-	-	-
	e i –	-	-	-	-	-	-
	e – l	-	1	-	-	1	-
	– i l	-	-	-	-	-	-
	e – –	-	3	-	-	9	-
	– i –	1	-	1	2	-	2
x	– – l	1	3	-	-	-	-
x	– – –	21	9	3	-	-	-
		totals			2	10	2
rule of usage		random selection			1/7	5/7	1/7
		selection by plurality			0	1	0

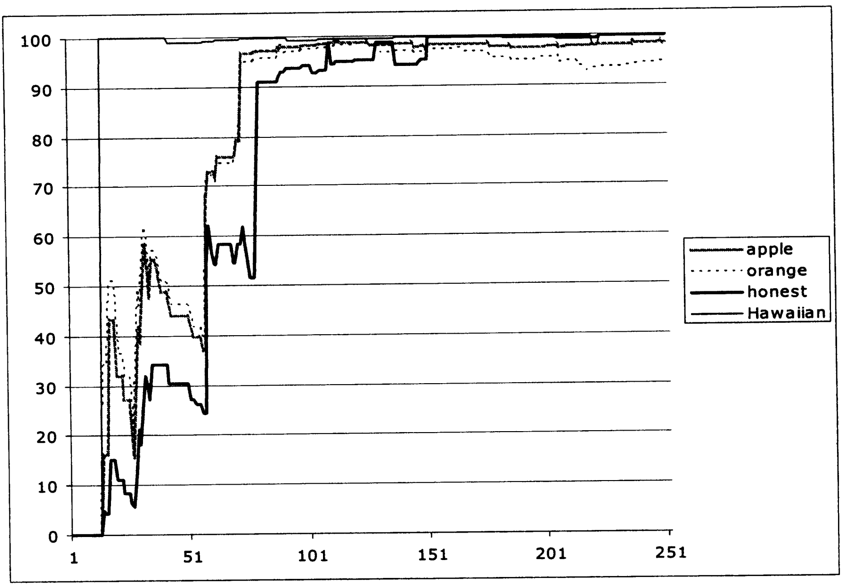
4. Applying analogical modeling to morphology

I begin this section by analyzing what is considered a classic categorical "rule", namely the rule for predicting the indefinite article *a*/*an* in English. As is well-known, the crucial variable is whether the following segment is a consonant or a vowel, thus *a boy* versus *an apple*.

When we treat this problem analogically, we are able to derive the typical rule-like properties that we expect, but other significant behavioral properties can be derived that no rule approach is able to predict or model. In the following example, the dataset for predicting *a* versus *an* is derived from two book chapters that I wrote over a year ago (and with no intention of using to construct this dataset). In all, there are 251 examples of the indefinite article in this dataset, of which 40 take *an* and 211 take *a*. Each *an* example is followed by a vowel-initial word, each *a* example by a consonant-initial word. So the data is "clean": the two chapters were correctly edited so that there are no exceptions to the "rule".

Our dataset is described by 9 variables based on the surrounding sounds, both preceding and following, which means that we include variables that we know in advance are not supposed to be "significant" or "crucial". We basically use the two preceding and the two following phonemes, along with their syllabic nature, and specify whether there is a phrasal break just before the indefinite article.

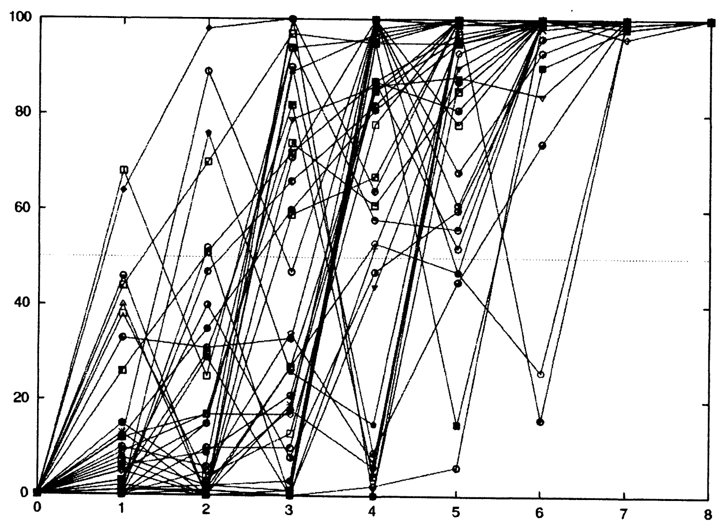
Using this dataset to predict the indefinite article for 8 different test items (not in the dataset), we get the following basic results for the full 251 items: (1) if the following word in the test item begins with a consonant, we get the "correct" *a* form virtually 100% of the time; (2) if the following word in the test item begins with a vowel, we get the "correct" *an* form about 98% of the time, but for 2% of the time we get leakage towards the "incorrect" *a*. This difference in leakage is quite dramatic when we consider a restricted dataset, starting out with only a few occurrences in the dataset and steadily increasing the number of occurrences. Initially, *a* is uniformly predicted no matter what kind of segment (consonant or vowel) follows, but as the data increases and tokens of *an* are introduced, the model begins to irregularly predict *an* when followed by a vowel, but over time moves inevitably towards systematically predicting *an*. Yet the model continues to predict some leakage towards *a* (but none towards *an* when followed by a consonant), even when adult-like behavior has been achieved (after about 80 occurrences of *a* and *an*). Below we have a graph of the predicted *an* behavior for the four vowel examples, with their one-way leakage decreasing erratically until stability is reached (but not with complete predictability). Note the small window of leakage that never fully closes.



On the other hand, the expected *a* is consistently predicted when the following segment is a consonant, even when the number of occurrences in the dataset is small.

We note from this example that AM is not simply reproducing the data. The data itself is fully "regular", having no exceptions to the "rule". The predicted fuzziness of AM justifies Sapir's well-expressed statement that "all grammars leak". And sometimes the leak is directional. Unlike rule approaches, there is no need to hunt for extra reasons to explain the performance. The system itself predicts the kinds of errors that occur, both for children learning the system and for adults who occasionally replace *an* by *a*. There is no need to set up a theory of markedness that would favor open rather than closed syllables (thus "explaining" why *an* tends to be replaced by *a*, but not vice versa). A variant explanation would be to declare that it is phonetically more difficult to pronounce *an boy* than *a boy*. This may be true enough for the youngest of children learning English, but it is not for adult speakers (or even moderately young children) who readily pronounce *one boy* as /wən bɔi/ rather than as /wə bɔi/.

The one-way leakage that we observe is inherent within the dataset itself and is due to the relative sparsity of *an* within the contextual space (when compared with *a*). This tendency can also be directly seen in the historical dynamics of the indefinite article. Using the same dataset for *a*/*an* as before (with 40 occurrences of *an* out of 251), we can model the historical drift in dialects of English towards replacing *an* completely by *a*. This particular model predicts that when there are less than about 70 accessible occurrences to predict from, there will be a steady S-transition from the original state of 40 occurrences of *an* to eventually none. Overall, the S-transition is very regular and smooth: the shifting starts out slowly with a few innovative cases, then the transition moves more rapidly, and in the end slows down, with a few relics holding out until the transition is complete. But when we consider the probabilistic predictions for each of the individual 40 cases of *an*, we see that the drift is highly erratic, with sharp shifts and frequent reversals in the predicted behavior, but eventually we get a complete shift from *an* to *a*.



Although the overall shifting seems quite tranquil, the specific examples of the drift are very turbulent. Actual historical examples of drift reflect this kind of turbulence.

5. Robustness

One important advantage of analogical modeling is that it can deal with what might be considered unexpected or defective language data such as dialectal developments and various errors found in adults' and children's language. One particular aspect of this robustness is the ability to make predictions even when the "crucial" variables are missing. For instance, in the case of the indefinite article *a/an*, suppose that the first segment in the immediately following word is overlaid with noise, yet we still wish to predict whether the article should be *a* or *an*. Using AM, we are still able to make predictions when the crucial variable is missing. Sometimes the resulting analogical set is dominated by a particular word, which basically means that the system has figured out what the partially obscured word is. Other times, no specific word may dominate the analogical set, but nonetheless redundancies in the dataset allow the correct prediction to be made. For instance, if the second segment in the following word is an *s*, then the analogical set will be dominated by examples for which the first segment (here obscured in the given context) is a vowel. This result derives from the fact that there are no consonant-*s* word-initial sequences in English; only vowel-*s* sequences are found. (Greek borrowings such as *psychology* are of course not pronounced with an initial /ps/ cluster in English.)

One important property of AM is its ability to predict leakage when given contexts are near exceptions. Sometimes unusual pronunciations of certain frequent words can affect nearby words. For many speakers of English (such as myself), the names for the days of the week have an alternative /i/ vowel pronunciation for the final *-day* morpheme. For instance, *Monday* has the pronunciation /məndi/ in addition to the standard /məndei/. This alternation leads speakers to accidentally produce the /ei/ pronunciation when /i/ is correct, but only in words that are close to the names of the days of the week, such as the burger establishment *Wendy's* being pronounced (by me) as /wɛndeiz/ or the example from a friend who first pronounced the name of the Utah community *Sandy* as /sændei/, then immediately corrected it to /sændi/. The nearby words *Wednesday* and *Sunday* serve as the source for the alternative /ei/ pronunciations, but only because there is the variant /i/ that allows this kind of phonological backformation.

Sometimes exceptional behavior is created because an item is extremely close to a single exemplar. Consider the word *consonantal,* which linguists typically pronounce as /kɑnsənɛntəl/ rather than the dictionary pronunciation /kɑnsənæntəl/. The obvious nearby analogy here is the word *continental.*

Other times a strong attraction occurs because of a whole gang of exemplars. Consider the pronunciation of the word *nuclear* as /nukjələ˞/ rather than the standard /nukliə˞/. As Jesse Sheidlower, an editor for the Oxford English Dictionary has pointed out (*New York Times*, 13 October 2002), the reason for the much-maligned innovative pronunciation is a very large gang of words that end in -/kjələ˞/, namely

particular, spectacular, circular, molecular, secular, perpendicular, muscular, *vernacular, jocular, vascular, curricular, ocular, corpuscular, follicular, specular, testicular, vehicular,* and *vesicular*. And many of these words are fairly frequent (the ones given in bold). On the other hand, there is only one word that supports the standard pronunciation for *nuclear:* namely *cochlear,* itself quite rare. Thus it is not surprising that speakers are tending to pronounce *nuclear* as /nukjəlɚ/, despite its spelling.

6. Not just nearest neighbors

Sometimes it has been argued that analogical modeling could be more efficient if it only looked for the nearest neighbor(s) to select an exemplar rather than considering exemplars that might be more distant. The following case involving the Finnish past tense clearly requires exemplars that are not the nearest neighbor. In trying to predict the past tense for the Finnish verb *sorta-* 'to oppress', there are two viable possibilities: *sorsi* and *sorti*. The first can be derived from the stem *sorta-* by replacing the *t* by an *s* and the final *a* vowel by *i*. We represent this outcome by the representation *tV-si*, where *V* stands for a vowel. The second possibility, *sorti*, can be derived by simply replacing the final *a* vowel by *i* and thus leaving the *t* of the stem unchanged. We represent this outcome as *V-i*.

Rule approaches to the Finnish language have predicted that the past tense for *sorta-* should be *sorsi*, which is the historical form. Yet actually *sorti* is the preferred form, and samplings show it clearly dominating over *sorsi*. We get the following evidence from the standard scholarly dictionaries of the language:

- *Nykysuomen Sanakirja* lists *sorta-* uniquely as its own morphological type which shows that *sorti* is clearly exceptional from a rule perspective. The alternative form *sorsi* is listed as a minor variant. Citations in the dictionary give 10 with *sorti*, only 1 with *sorsi.*
- *Suomen Kielen Perussanakirja* gives *sorti* as the imperfect (that is, past tense), while *sorsi* is listed as rare. There are 2 citations with *sorti*, none with *sorsi*.

Occurrences of *sorsi* are found, such as in the 10 June 2001 issue of the *Helsingin Sanomat*, the major daily newspaper in Finland, which had the following heading to a brief article: "Omat naiset sorsivat Lipposta", which means, 'Own women oppressed Lipponen'. (Lipponen was at the time the prime minister in the government.)

Now the question is: How can we explain this preference for *sorti* over *sorsi* (the historical form), especially when all rule approaches predict *sorsi* for the standard language? One possibility is that the nearest neighbor makes the right prediction, but unfortunately the nearest neighbor is *murta-* 'to break', which takes the *tV-si* outcome, namely *mursi*. In other words, the nearest neighbor predicts *sorsi*.

When we examine the analogical set for the verb *sorta-*, we discover that the main reason for *sorti* is the *o* vowel, a factor which no analyst (Finnish or otherwise) had ever come up with prior to the AM analysis of the Finnish past tense. When we consider the analogical set, we see that the nearest gang, representing verbs ending in *rtA* (where *A* is a low vowel, either *a* or *ä*), is rather minor in its ability to provide exemplars. On the other hand, there is a huge gang, but further away, that has *o* as the first stressed vowel. Every one of these verbs takes the *V-i* outcome, thus overall *sorti* is predicted about 94% of the time. (See the schematic drawing of the analogical set for *sorta-* in Skousen, Lonsdale, and Parkinson 2002:33.)

One very important result from this example of *sorta-* is that we must not assume in advance which variables are the crucial ones and thus ignore all the others. The *o* vowel is definitely not a significant variable in the historical development of the Finnish past tense. And for all other verbs in the language, the *o* vowel is not the crucial factor. Its potential use remains latent until the appropriate given context is chosen.

7. The rule equivalent

Analogical modeling can be re-interpreted in terms of rules, as follows: (1) every possible "true" rule exists; and (2) the probability of using a "true" rule is proportional to its frequency squared. By a "true" rule, I mean that the context for the rule is homogeneous in behavior. Homogeneity occurs under two conditions: either the rule is deterministic in behavior (has only one outcome); or if the rule is non-deterministic, no subrule of the rule can behave differently.

Despite this equivalence, AM is not like regular rule approaches. First of all, there is no partitioning of the contextual space. Since all (!) the "true" rules are said to exist, there will be overlapping rules, redundant rules, and rules based on as little as one occurrence. Secondly, these equivalent "true" rules, when considered from the perspective of AM, are created "on the fly"; they are not stored somewhere, waiting to be used.

A third, and most crucial, difference is how AM treats non-deterministic rules (that is, rules that are probabilistic in nature and sometimes referred to as "variable rules"). Consider a hypothetical probabilistic rule for the past tense of the English verb *dive*. Let us assume that under some specific conditions, the probability of producing the irregular *dove* as the past tense form is 2/3, while the probability of the regular *dived* is 1/3. One immediate problem that arises in dealing with probabilistic rules is whether the context for this rule will ever be homogeneous. It seems very unreasonable to assume that every subcontext of this probabilistic context will have precisely the same probability distribution of (2/3,1/3). In fact, it also seems very doubtful that the probability would be a nice clean ratio like (2/3,1/3). Real coins, for instance, do not have a precisely equal probability of (1/2,1/2) for coming up heads and tails. Rather, actual occurrences suggest that the ultimate objective probability for a coin coming up heads is some irrational number near 1/2, not 1/2 itself. Perhaps after many trials on a particular coin, this probability can be estimated by a rational number (say something like 0.498157). Yet even this estimate is unstable since the coin itself would be affected by all that flipping and thus the actual probability of heads would change as the coin continued to be physically tested.

The most serious issue with probabilistic rules is how to actually learn a probability and then use that probability to predict behavior. In the case of *dove/dived*, the learner would be confronted with a finite sequence of past-tense forms:

o o d d o d d o o o d o o o d d o o d o o o d o d o d o o d d d o d

(Here *o* stands for *dove* and *d* for *dived.*) From such finite sequences, a probability must be determined. One could use the sample ratio, which in this case would be (19/34,15/34). In this instance, I used a random number table to generate this sequence, and although my linguistic rule may have an objective probability of (2/3,1/3), its sample ratio is hardly ever equal to that probability. The task of determining the objective probability is very hard indeed. But the really hard part is using an objective probability, even if it is the sample ratio, to predict probabilistic behavior randomly. The random sequences used in computer programs are based on mathematical functions that look random enough over finite intervals, but ultimately involve non-randomness.

It was this major problem of learning and using probabilities that led me in 1979 to realize that the solution was much different. Instead of dealing with probabilities, it was much simpler to directly store the examples, say of *dove* versus *dived,* and then, when there was a need to predict the past tense of *dive*, to randomly select one of those previous examples, examine whether it was *dove* or *dived*, and base the prediction on that single randomly-selected example. Thus the original motivation for analogical modeling was to solve the problem of probabilistic rules by refusing to posit them! Instead, we store and use examples. In this case, about two thirds of the examples will be *dove*, about one third *dived.* In fact, there will be no objective probabilities at all, only the occurrences.

8. Quantum computing of analogical modeling

8.1 The exponential explosion

One well-known property of analogical modeling is the exponential explosion: whenever we add one variable to the system, we basically double the running time and the memory requirements needed to determine the analogical set for a given context. This exponentiality results in pragmatic difficulties for AM since no matter how powerful the computer being used to determine the analogical set, there is a restriction on the number of variables that can be used for any one problem. The original AM program (see Skousen 1989) was restricted to only 12 variables. Subsequent work on the algorithm now allows us to posit up to about 30 variables, but there is still (and will always be) a fairly low bound on the number of variables that can be implemented on a classical computer. Various pragmatic approaches to resolving this problem have resulted in some reduction in the exponentiality, but none have been able to eliminate it.

Within the last few years, however, I have discovered striking parallels between AM and quantum mechanics, in particular quantum computing, which suggest that AM can be naturally interpreted as a variant of quantum processing. Quantum computing allows the simultaneous processing of 2^n states instead of a single state (as in a standard sequential computer). I will outline here some of these parallels.

8.2 Superpositioning, interference, and reversibility

Quantum mechanics permits n quantum bits (called qubits) to simultaneously represent 2^n states and to evolve through time by means of reversible operators. This reversibility prevents any loss of information while the 2^n states exist in this superposition of states. Such a superposition permits the multiple existence of the same "object". As the system evolves, objects can interfere and can become entangled. Further, the chances that a given state will be ultimately observed is increased or decreased. Objects that behave alike (or are "in phase") constructively interfere, while objects that behave differently (or are "out of phase") destructively interfere.

Analogical modeling follows the same basic process. Given n variables for a given context, 2^n supracontexts are defined, one supracontext for every possible combination of the n variables. As data occurrences are read in and simultaneously assigned to all the applicable supracontexts, the homogeneity of each supracontext is determined by means of reversible operators. The condition of supracontextual homogeneity prevents any increase in disagreement, thus no loss of information occurs. The same data occurrence (or "object") is typically found in more than one homogeneous supracontext, with the result that homogeneity leads to proximity and gang effects. In contrast to the constructive interference of homogeneity, all heterogeneous supracontexts are zeroed out (a kind of destructive interference).

8.3 Observation and the squaring of the underlying linearity

In quantum mechanics, the probability of each state being observed is not directly represented. Instead, a linearly defined amplitude (usually a complex number) is assigned to each state. The superposition of states and its parallel processing continues until observation (or measurement) occurs. Then the superposition collapses into a single state and the probability of that state occurring is equal to its amplitude squared.

In analogical modeling, each homogeneous supracontext is linearly represented by the number of data occurrences assigned to that supracontext. (Heterogeneous supracontexts are assigned an amplitude of zero, which means that none of their data occurrences can be accessed.) In a homogeneous supracontext, pointers are assigned between each pair of occurrences. Data input continues until the decision is made to observe the analogical set, yet when that occurs, only one of the homogeneous supracontexts is accessed. Moreover, we select a pointer to an occurrence, not an occurrence itself, with the result that the underlying linear measure based on the number of occurrences is squared.

8.4 Exemplar-based quantum computing

One interesting aspect of quantum computing is that reversibility requires that input data can never be erased, from which we may conclude that quantum computing itself is an exemplar-based system and that quantum computation of any language-based system will be an exemplar-based one. Further, analogical modeling (using random selection) appears to be a general quantum computing algorithm. The exponential explosion in AM is not inherent, but instead is the result of using classical computation.

References

Skousen, Royal. 1989. *Analogical Modeling of Language.* Dordrecht: Kluwer.
Skousen, Royal. 1992. *Analogy and Structure.* Dordrecht: Kluwer.
Skousen, Royal, Deryle Lonsdale, and Dilworth S. Parkinson. 2002. *Analogical Modeling: An exemplar-based approach to language*. Amsterdam: Benjamins.

Department of Linguistics and English Language
Brigham Young University
Provo, Utah 84602

royal_skousen@byu.edu

A Common Structure for Cross-linguistic Conjunction Patterns

MARCUS SMITH
University of California, Lost Angeles

0. Introduction

In this paper, I investigate a series of patterns and asymmetries found in coordinate structures cross-linguistically and propose a simple structure that can accommodate the observed data. To keep the task manageable, I will limit the investigation to DP conjunctions which can (or must) be interpreted as all members of the conjoined expression being involved in the same action; that is, (1a) rather than (1b).

(1a) John and Mary went to the market together.
(1b) John and Mary went to the market separately.

I will abstract away from some otherwise important features of coordinate structures, mainly because they do not seem to be relevant to the issue under question. The linear order of conjuncts and the conjunction will not be addressed. Also, I will not address the issue of what morphological category the "conjunction" is in far a given language.

1. Basic Data

Four languages suffice to show the basic range of coordinate structures we shall account for. The languages used were selected to avoid genetic and areal bias: each is from a different language family, and each was originally spoken on a different continent.

English (Indo-European) has a simple particle serving as the conjunction and at least two conjuncts (2). Since this pattern is the most common cross-linguistically, most theories of coordination have assumed it to be basic.

(2) I saw **John and him**.

Pima (Uto-Aztecan) is similar to English, in that it uses a conjoining particle, and contains two or more conjuncts. However, the two languages differ in that

Pima (non-subject) conjunctions are associated with a pronominal clitic that attaches to the licensing head. The clitic has the person and number features one might expect of a pronoun that would replace the conjoined expression. For example, in (3) the clitic is first person plural, because the conjoined expression includes a first person pronoun and two individuals (Smith in progress). Similar structures are found in the closely related language Tohono O'odham. (Zepeda 1983) Borrowing terminology from Lichtenberk (2000), I will refer to this sort of pronominal element as an **inclusory pronoun**.

(3)	John	'at	**t-**	naam	**'aañi**	**c**	**heg**	**Mary**.
	John	AUX:PF	1PL-	meet	1SG	and	DET	Mary

'John met me and Mary.'

With only slight changes to the Pima structure, we get those found in Mapuche (Araucanian). (Smeets 1989) This language also has inclusory pronouns, though these are not clitics. Rather, they occur as a constituent with the conjoined expression. As with Pima, the inclusory pronoun reflects the person and number of the entire conjoined expression. Unlike English and Pima, the pronominal conjunct highest in animacy is optional (4a *vs.* b).

(4a)	**Iñché**	**eymi**	**iñchiu**	iyu.
	1SG	2SG	1DU	eat:1DU

'You and I ate.' (Smeets 1989:177)

(4a)	**Eymi**	**iñchiu**	iyu.
	2SG	1DU	eat:1DU

'You and I ate.' (ibid.)

The "special coordinate construction" in Tagalog (Austonesian) is very similar to the Mapuche pattern in (4b). There is typically one plural pronoun, inflected to show the person and number of the entire conjoined expression, and a set of at least one accompanying DP. Unlike Mapuche, a conjunct corresponding to the person of the inclusory pronoun is consistently absent (5). (Schachter and Otanes 1972)

(5)	Nakita	ko	**sila**	**ni**	**Juan.**
	saw	1SG	3PL	DET	Juan

'I saw him and Juan.' (Schachter and Otanes 1972:116)

2. Structural Generalizations

2.1 A Common Structural Basis

English and Tagalog initially appear to be quite different structurally. English has, but Tagalog lacks, a conjoining particle. Tagalog has, but English lacks, an inclusory pronoun. Tagalog lacks of simple singular pronoun in (5), and English cannot do without one in the same contexts. The only element they both have is at least one DP conjunct.

Looking at the other two languages, however, shows that there is a gradation from English-like to Tagalog-like patterns. Pima added an inclusory pronominal to an English-like structure, and Mapuche in turn removed the conjunction and optionally the highest animate conjunct from a Pima-like structure. Thus, English structures can be transformed into Tagalog structures (and vice versa) through a series of discreet changes with no significant gaps in the intervening structural types.

To accommodate the degree of variation seen, we need structures minimally like (6) (disregarding the linear order). Individual portions of this basic structure can be omitted to achieve the pattern found in a particular language.

(6) $[_{DP}$ Pronoun $[_{DP}$ X & Y]]

2.2 Lack of Inclusory Pronouns

Which parts of this structure may be omitted is not random: there are only three basic possibilities, though they can be combined. First, it is quite common for languages of the world to lack a clear example of an inclusory pronoun; most the European languages fit this description. The question this raises is whether the inclusory pronoun is lacking from these languages altogether, or whether it is simple covert.

The answer may depend on the assumptions of the framework one is working in, but it is not unreasonable to believe that such pronouns are present universally. In order to express an inclusory pronoun, the speaker must calculate the person and number of the entire conjoined expression. The same process can be seen in subject agreement in, *e.g.* English and Spanish. Thus, while English may lack an overt inclusory pronoun, it does invoke a comparable mental computation. If one views inclusory pronouns merely as the overt manifestation of a universal computation, then representing this computation in one's framework in a manner attested in multiple languages is a reasonable step.

2.3 Lack of Conjunctions

Some languages lack on overt coordinator, as we saw with Mapuche and Tagalog. The same question arises: Are the coordinators covertly present or completely lacking?

Johannessen (1994) gives diachronic evidence suggestive that coordinators are always represented, even if not expressed. She notes that languages have been known to borrow conjunctions, and doing so does not appear to affect the syntactic structures of these languages. The conjunction is cleanly assimilated. She suggests that if the addition of such a morpheme were an actual change in the structure, rather than mere adding phonological substance to an abstract concept, we would expect to see other changes, perhaps in case, agreement, or word order possibilities.

Another argument in favor of silent conjunctions comes from the possible semantics of the constructions. In most languages, in most contexts, juxtaposition results in a conjunction, as in the Mapuche examples above; but occasionally it can result in a disjunction. The constructions in English (7) present alternative figures as possible prices. In such contexts, it would also be pragmatically reasonable to have a "respectively" reading, where different participants paid different amounts; but this interpretation is not available.

(7) We paid six, seven dollars for the tickets.

The interpretation of juxtaposition, then, is conventionalized in particular languages or constructions. This idiosyncrasy must be represented somehow in the structure of the language. Similar to inclusory pronouns, if a conjunction is simply an overt manifestation of a mental process that is universally present, then it is reasonable to encode this process in a manner attested in other languages.

2.4 Missing Conjuncts

The Mapuche examples show that some language can optionally lack an individual conjunct. Assuming the structure suggested in (6), the special coordinate structure of Tagalog is required to be missing a conjunct. The individuals that would have been represented by an overt conjunct are semantically present in the sentence and accessible to the morphosyntax, as seen by the person and number of the inclusory pronouns and verb agreement.

The choice of which conjunct can be missing is tightly constrained. Typically, the missing conjunct must the member of the conjoined expression that is highest on the standard animacy heirarchy (8). In Mapuche example (4b), the missing conjunct is the first person pronoun rather than the second person pronoun.

(8) speaker > addressee > 3rd person pronoun > noun

Though it is difficult to tell in many cases, it appears that the missing conjunct is always a peripheral conjunct, never one between other conjuncts. Evidence of the position of a missing conjunct can be found in languages like Modern Irish (Indo-European), where only the first conjunct can be missing. In (9), a pronoun denoting the speaker is not present, but the contrastive particle *–se* that should attach to it is present, showing us where the pronoun should be.

(9)	Chaithfinn	**-se**	**agus**	**mo**	**chuid**	**fear**	muscailt.
	must(COND:1SG)	-CONTR	and	my	share	men	wake.up

'I and my men would have to wake up.' (McCloskey and Hale 1984:501)

A missing initial conjunct was also possible in Old Irish, though the exact details differ a bit from Modern Irish. The sentence in (10) has a preposition inflected for second person plural. It is immediately followed by a conjunction and a name, no second person pronoun is present. Old Irish conjunctions, like those in Modern Irish, typically occur between the conjuncts, showing that the initial conjunct is the one missing.

(10) comrac dúib **ocus** **Chú-Chulainn**
encounter between(2PL) and Chú-Chulainn(NOM)
'an encounter between you (sg) and Chu-Chulainn' (Thurneysen 1998:156)

The pattern as to which conjunct can be missing overlaps with other asymmetries in coordinate structures. In the remainder of this paper, I will illustrate some of these asymmetries and alter the structure in (6) to capture the generalizations found. I will refer to the conjunct that can have the properties discussed below as the **primary conjunct**.

3. The Primary Conjunct

The primary conjunct can be distinguished from other conjuncts through its syntactic behavior. It can have specific requirements imposed upon it not found on other conjuncts, and often this conjunct interacts with the rest of the sentence in manners not permitted of the others. For example, I showed above that if a language allows null conjuncts, it conform to a strict set of conditions, with the result that only one is possible in a given expression. This is the sort of specific requirement intended. Three other properties will be inspected here: a restriction on conjunct ordering, agreement, and case marking. The behaviors converge in suggesting that the primary conjunct is a syntactically determined position.

3.1 Conjunct Ordering

I mentioned above that when it can be determined, missing conjuncts tend to be peripheral elements. They also conform to the animacy hierarchy. In some languages, the grammatical ordering of conjuncts seems to follow similar tendencies.

A good example of this is Ulithian (Austronesian). In this language, whenever a pronoun is coordinated with a noun, the pronoun must occur first. Thus, in (11a) the order pronoun-noun is grammatical, but in (11b) the order noun-pronoun makes the sentence unacceptable. Because there is a syntactic requirement placed on the first conjunct in relation to other conjuncts, it should be identified as the primary conjunct for this language.

(11a) **Gaag** **mė** **Coon** melee xa sa koxo.
I and John FM PM PERF go
'I and John went.' (Sohn and Bender 1973:209)

(11b)* **Coon** **mė** **gaag** melee xa sa koxo.
John and I FM PM PERF go
'John and I went.' (ibid.)

The English requirement that a nominative first person pronoun be the final conjunct in a coordinated subject is rather curious, and probably should not be used as evidence in determining the primary conjunct of English. English *I* does not pattern like typical conjuncts; in particular, its case seems to be dependent on the position in the conjoined expression, rather than its grammatical role, as in (12).

(12) All debts are cleared between **you and I**. (Shakespeare, *Merchant of Venice*)

While "improper" cases frequently occur in natural language conjunctions (as discussed in section 3.3 below), the pattern seen here is not expected of a primary conjunct. Typically, this element gets marked with the case one would naively expect, and the other conjuncts vary in surprising ways. I will not address the issue in detail, but refer the reader to previous work on the matter, such as Emonds (1986).

3.2 Verbal Agreement

In some languages, agreement between a verb and a conjoined expression can hold either with the entire expression or with just a single conjunct. As with the condition on missing conjuncts, this kind of asymmetric agreement pattern always holds with a peripheral conjunct.

One example of this has already been presented, the Modern Irish sentence in (9). The verb is inflected for a first person singular subject, though the syntactic subject is a conjunction. Significantly, the same sentence has a missing primary conjunct, and that missing element is the very one the verb shows agreement with.

Mapuche also has this sort of construction, illustrated in (13). The first example has a verb showing first person dual agreement, as one might expect. The second example shows asymmetric agreement. The verb is inflected for first person singular, though the inclusory pronoun shows that the subject is a conjunction. As with the Modern Irish example, the asymmetric agreement holds with a pronoun that is not overtly present in the structure, showing an overlap with the missing conjunct pattern.

(13a)	fey	iñchiu	nütram	-ka	-y	-u
	he	1DU	conversation	-FAC	-1NS	-DU
	'We talked together.' (Smeets 1989:178)					
(13b)	fey	iñchiu	nütram	-ka	-n	
	he	1DU	conversation	-FAC	-IND:1SG	
	'I talked with him.' (ibid.)					

This phenomenon has been discussed extensively in the literature, and the reader is referred to that work for detailed discussion, especially Corbett (1991, 2000), Munn (1993), and Johannessen (1998).

3.3 Case Marking

A less commonly discussed form of asymmetric agreement pattern is the case marking of individual conjuncts. The expected situation is for all the conjuncts to be case marked according to the grammatical role of the entire expression. It is not uncommon, however, for only a single conjunct to receive the proper case, and for the remaining conjuncts to receive a default case.

Kiparsky (1968) discusses this in relation to the ancient Indo-European languages Old Irish, Ancient Greek, and Sanskrit. All three show the same basic pattern, though only Old Irish will be illustrated here (14). In each of these languages, the first conjunct is case marked according to the grammatical role of the entire expression, but each subsequent conjunct is (or may be) marked with nominative case.

(14) rí dorigni aéar n-úar ocus tene réil
king has.made air cold(ACC) and fire clear
rorúad ocus talam bladmar brass
very.red(NOM) and earth glorious great(NOM)
'The King has made the cold air, and the clear red fire, and the glorious great earth.' (Kiparsky 1968:54)

Asymmetric case assignment also occurred in prepositional phrases, as shown in (10), repeated here. The second conjunct is in the default nominative case and the conjunct that should receive the typical accusative is missing.

(10) comrac dúib ocus Chú-Chulainn
encounter between(2PL) and Chú-Chulainn(NOM)
'an encounter between you (sg) and Chu-Chulainn' (Thurneysen 1998:156)

(See also Johannessen (1998) for discussion of asymmetric case assignment.)

3.4 Adjusting the Structure

Various asymmetric properties of coordinate structures have been shown to correlate with each other in syntactic structures. The conjunct that can be missing tends to be the same as the conjunct involved in asymmetric agreement and asymmetric case assignment. The congruence of all these properties suggests that the primary conjunct is a syntactic notion.

The interesting thing is that the properties of the primary conjunct are typical properties of arguments, whereas the properties of the other conjuncts are deviant. In many languages, arguments are optionally present (such as the null subjects in Romance) while adjuncts do not have the same freedom. Agreement on a verb or adposition is typically triggered by an argument. Similarly, primary conjuncts are case marked as an argument would be. These facts suggest that the primary conjunct and not the others is the main argument of the expression.

To capture this, I propose that the coordinate structure suggested in (6) should be modified to that in (15). The primary conjunct here is the X and is the head of the structure. The conjunction and the Y conjunct is adjoined to the primary conjunct..

(15) $[_{DP}$ Pronoun $[_{DP}$ X $[_{COP}$ & Y]]]

Except for the inclusory pronoun, the structure above is essentially that proposed by Munn (1993) and Mel'cuk (to appear) based on other considerations. They point out that this structure has other beneficial properties. One very beneficial property relates to selection. Coordinated DPs have the same distribution as simple DPs. By placing one conjunct as the head, and having the other conjuncts adjoined to it, this distributional fact is captured without any complex, construction specific rules.

3.5 Secondary Conjuncts

An issue remains about the behavior of the adjoined conjunction phrases. Why is it that sometimes they behave like regular arguments, receiving the proper case and contributing to verbal agreement, but other times behave as though isolated from the rest of the structure?

I suggest this is a consequence of the dual nature of these structures. The syntax of these constructions seems to be asymmetric, but the semantics is symmetric: all the conjuncts participate equally in the event described. From a syntactic perspective, only the primary conjunct should behave as typical of arguments. From a semantic perspective, they should all behave the same. This discrepancy results in *a priori* unexpected patterns, where for some portions of the grammar the rules follow syntax but semantics for others.

This predicts that optional asymmetries should correlate with semantic differences. There is weak evidence that this is true. In the Mapuche example (13), repeated here, optionality of asymmetric agreement is correlated with a difference in translation. The symmetric version is translated in a way suggesting both participants are equal, while the asymmetric version seems to place more emphasis on the (silent) individual controlling the agreement.

(13a)	fey	iñchiu	nütram	-ka	-y	-u
	he	1DU	conversation	-FAC	-1NS	-DU
	'We talked together.' (Smeets 1989:178)					
(13b)	fey	iñchiu	nütram	-ka	-n	
	he	1DU	conversation	-FAC	-IND:1SG	
	'I talked with him.' (ibid.)					

Further evidence can be seen in Arabic, where asymmetric agreement correlates with the distribution of the participants in the event. When the verb has full agreement with a conjoined subject, the sentence can be interpreted as having all the individuals together in the event (16a). When agreement is asymmetric, this interpretation is no longer acceptable (16b). (16c) shows that asymmetric agreement is available generally.

(16a) Raacrho Kariim w Marwaan sawa.
left.PL Kareem and Marwaan together
'Kareem and Marwaan left together.' (Aoun, Benmamoun, and Sportiche 1994)

(16b)* Raacrh Kariim w Marwaan sawa.
left.3SG.MASC Kareem and Marwaan together
'Kareem and Marwaan left together.' (*ibid.*)

(16c) Gatal ?el- walad we- l- banaat ?el bisse
killed.3SG.MASC the- boy and- the- girls the cat
'The boy and the girls killed the cat.' (van Oirsouw 1987:232)

4 Conclusion

Despite initial impressions that conjoined expressions across languages can be quite distinct, an appropriate sample of languages shows that the differences are confined to a small range of possibilities. Asymmetries in case assignment, agreement triggering, and null conjuncts converge on a simple pattern suggestive that only a single conjunct acts as the head of the structure and others are adjoined to it.

References

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. Agreement, word order, and conjunction in some varieties of Arabic. *Linguistic Inquiry* 25:195-220.

Corbett, Greville. 1991. *Gender*. London: Cambridge University Press.

—. 2000. *Number*. London: Cambridge University Press.

Emonds, Joseph. 1986. Grammatically deviant prestige constructions. In *Festschrift for Sol Saporta*, ed. Michael Brame, Heles Contreras, and Frederick Newmeyer, 93-129. Seattle: Noit Amrofer.

Johannessen, Janne. 1998. *Coordination*. New York: Oxford University Press.

Kiparsky, Paul. 1968. "Tense and mood in Indo-European syntax" Foundations of Language, 4(1):30-57.

Lichtenberk, Frantisek. 2000. Inclusory Pronominals. *Oceanic Linguistics* 39: 30-57.

McCloskey, James and Kenneth Hale. 1984. On the Syntax of Person-Number Inflection in Modern Irish. *Natural Language & Linguistics Theory* 4: 487-533.

Mel'cuk, Igor. To appear. Dependency in Linguistic Description. Ms.

Munn, Alan. 1993. *Topics in the Syntax and Semantics of Coordinate Structures*.

Ph.D. thesis, University of Maryland.
Schachter, Paul and Fe Otanes. 1972. *Tagalog Reference Grammar*. Los Angeles: University of California Press.
Smeets, Ineke. 1989. *A Mapuche Grammar*. PhD Thesis, Rijksuniversiteit te Leiden.
Smith, Marcus. In progress. Partial Agreement in Pima. MA thesis, UCLA.
Sohn, Ho-Min and B. W. Bender. 1973. *A Ulithian Grammar*. Canberra: The Australian National University.
Thurneysen, Rudolf. 1998. A Grammar of Old Irish. Dublin: Dublin Institute for Advanced Studies. Translated by D. A. Binchy and Osborn Bergin.
van Oirsouw, Robert. 1987. *The Syntax of Coordination*. New York: Croom Helm.
Zepeda, Ofelia. 1983. A Papago Gramar. Tuscon: University of Arizona Press.

Marcus Smith
Department of Linguistics
3125 Campbell Hall
Los Angeles, CA 90095-1543

smithma@ucla.edu

Indexicals in Dialogic Interaction[1]

MAKIKO TAKEKURO
University of California, Berkeley

0. Introduction

We often find in interaction an array of linguistic elements in one speaker's utterance that make use of elements in another's. This phenomenon -people making use of other people's utterances- is called "dialogicality". Du Bois (2003) explains "dialogicality", based on Jakobson's (1966) notion of "parallelism" and Bakhtin's 1981[1935] notion of "voice", as "my words come from and engage with your words, and with the words of those who have spoken before us". In interaction, dialogicality is pervasive between comparable linguistic elements at any level: words, morphemes, phonemes, features, meanings, referents, illocutionary forces, etc.

Instances of linguistic forms such as honorifics and sentence-final particles can be dialogically interconnected to one another in interaction. But analysis focused on dialogicality of these linguistic forms has been scarce. Taking a combined interactional and semiotic approach to the analysis of honorifics and sentence-final particles in Japanese, I examine the ways in which these linguistic forms construct dialogic interaction. Seeing linguistic forms as indexical signs, which will be explained in the next section, I point out dynamic changes in the indexicals in dialogic interaction. I also demonstrate that the use of the linguistic forms interacts with the situated interactive context and larger social norms. Lastly, I discuss socially significant meanings of dialogic interaction within the framework of social interaction, referring to Goffman's (1956) notions of "deference" and "demeanor". Using these notions to interpret the meanings of dialogic interaction, I elaborate on what the speaker owes to others and to oneself

[1] An earlier version of this paper in part appeared in Takekuro (2002) and was also presented at the Workshop on East Asian Languages (WEAL) at University of California, Santa Barbara in March 2003. I wish to thank the participants in the 2003 meeting of the Berkeley Linguistics Society and WEAL for their questions and insightful comments. This paper significantly benefited from discussions with Bill Hanks. I am deeply grateful to Robin T. Lakoff who heard, read, and criticized various versions of this research. Responsibility for the data and analysis rests solely with me.

in social interaction. Through dialogic interaction, the speaker displays ordered activated affinities and the sense of self in meaning and structure.

0.1. Indexicality

An "index" is one of the constituents along with "icon" and "symbol" in Peirce's (1955) tripartite system of signs. In Peirce's terminology, an index is representative of the object by virtue of "being really affected" through a dynamic or causal relation to the object. For instance, a knock on the door is an index of the presence of a visitor; a weathercock is an index of the direction of the wind. When a sign is an index, it stands for the object neither by similarity nor convention, but by contiguity with it. In this sense, an indexical sign is existentially bound to the object.

The adjective "indexical" and noun "indexicality" are used to describe linguistic signs that signal or point to certain features of the communicative context (Jakobson 1960, Lyons 1977, Morris 1938, Peirce 1955, Silverstein 1976). Linguistic indexicals include regional accent, pronouns, demonstratives, deixis, tense, and honorifics, whose tokens stand in dynamic and existential relations to their objects. As indexicals bear a direct connection with the object, the interpretation of indexical signs depends on the context in which it occurs.

Similar notions of indexicality are worth noting. Gumperz (1982) has identified a subclass of indexical signs, which he calls "contextualization cues". Contextualization cues indicate how an utterance is to be understood and what its rhetorical role in a sequential discourse is, therefore invoke the framework of interpretation of sociocultural context. Goffman (1974) defines "footing" as the position or alignment an individual takes in uttering a given linguistic expression. Bakhtin (1981[1935]) presents the notion of "voice". In interaction, participants use various "voices", such as reporting someone else's speech, mimicking someone, and speaking as someone else, all of which are indexed by linguistic features. A variety of these notions describing the more or less the same phenomena of indexicality suggest that the indexical function of language is central to communicative practice and serves to establish social relationships in context.

Silverstein's (1976) view of indexicality is most relevant to this study. Following semiotic traditions of Peirce, Jakobson, and Jespersen, Silverstein presents a two-way classification of indexical types: presupposing and creative. A presupposing indexical sign points to some contextual aspect independently known. In this sense, the sign presupposes the aspect. A creative indexical sign can make a particular contextual feature operative in the communicative context, by picking out the referent. For example, an honorific expression such as *vous*, on the one hand, functions as a presupposing index when it points to the addressee's higher status in a social context where status difference exists between interlocutors. On the other hand, the use of *vous* to a friend who is commonly referred to by *tu* can function as a creative index when it foregrounds relevant aspect of the context, such as deference, coldness, irony, humor, or sarcasm. Because of these two aspects, indexicals become primary tools to maintain and create social and psychological worlds among interlocutors.

In this paper, I take this semiotic notion of indexicality as the point of departure for analyzing honorifics and sentence-final particles in informal conversations among Japanese speakers in their 20s and 30s. I use the terms, "unmarked" and "marked" as analytical tools.[2] My aim was to determine how linguistic forms communicate indexically in the processes of interaction and interpretation.

0.2. Honorifics and sentence-final particles in Japanese

In the present study, the focus of honorifics is on two major categories: referent and addressee honorifics. Referent honorifics mostly appear as pronouns and predicate forms and have two further subclasses: *sonkei-go* ('respectful form' or subject honorifics) and *kenjo-go* ('humble form' or object (non-subject) honorifics). They commonly refer to the subject or non-subject that is socially distant from the speaker. Addressee honorifics, the so-called *teinei-go* ('polite form'), appear as verbal suffixes or as suppletive forms of the copula, generally used as markers of formality or in reference to an addressee who is socially distant from the speaker.

Sentence-final particles (hereafter, SFPs) in Japanese have traditionally three gendered categorizations: "feminine"; "masculine"; and "neutral" (Okamoto & Sato 1992, Okamoto 1997). Feminine and masculine forms have two subclasses: "strongly feminine/masculine" and "moderately feminine/masculine". Feminine forms are conventionally used by women, masculine forms by men, and neutral forms by both women and men.

Both honorifics and SFPs are highly indexical in the sense that they have pervasive context-dependency. Honorifics are forms of speech that signal deference, through conventionalized understandings of some aspects of the form-meaning relationship. They imply and grammatically encode participant role structures, relating the giver of deference to its recipient. SFPs are forms of speech that signal not only the speaker's gender but also the speaker's epistemic and affective stances and other relevant aspects of the interactive context.

1. Method

The database for this study consists of three spontaneous conversations among friends: a 120-minute conversation among ten females, ages 21 to 23; a

[2] The concept of markedness and its definitions have been much discussed in the literature. In the context of the present work, I will be using both "marked" and "unmarked" in a simple, restricted sense. "Marked" in this paper will refer to a linguistic form that is statistically more likely to occur, while "unmarked" will apply to a form with a rare statistical likelihood of occurrence. I am using this terminology in order to avoid the use of confusing terms such as "semantic presupposition", although Silverstein's concept of "presupposing" does not mean semantic presupposition.

120-minute conversation among seven females, ages 26 to 28; and a 120-minute conversation among two males and five females, ages 28 to 36. All the subjects speak more or less standard Japanese and resided in the eastern part of Japan (Kanto) at the time of the recordings, although three subjects in Data Set III grew up in other regions in Japan, such as Hokkaido, Tokushima, and Nagoya.

Before turning to the analysis of the data, I present a brief summary of how indexicality operates in Japanese informal peer group conversations. Statistically speaking, in an informal peer group conversation, young female speakers of Japanese in their 20s use honorific forms in no more than 0.70-1.92% of all intonation units (Takekuro 2002). Studies based on different databases (Okamoto 1997, Takekuro 2002) show that the most frequently used SFPs by female speakers of Japanese in their 20s are gender-neutral forms (68-70%), and the next frequently used ones are moderately masculine forms (18-19%) and moderately feminine forms (8-10%). They rarely use strongly feminine forms (0-5%) or strongly masculine forms (0-1%). Sturtz (2001) demonstrates that the most frequently used SFPs by male speakers in any generation are gender-neutral forms and the next most commonly used ones are moderately masculine forms. Due to the low frequency in usage, if honorifics or strongly gendered forms of SFPs are used in informal peer group conversations, such linguistic forms could be viewed as marked. However, in this study, I will demonstrate that many of the seemingly "marked" use of these linguistic forms can be reinterpreted as "unmarked" in situated interactive contexts. Indexicality of linguistic forms is dynamic and constantly negotiated in situated contexts.

2. Data analysis

In this section, I will present examples of dialogic interaction in the use of honorifics and SFPs in Japanese informal conversations. Dialogic interaction is influenced by conversational topic and becomes most prominent in the context of a conversational topic shift.

2.1. Honorifics

When one speaker uses an honorific form, although it is not a norm in informal peer group conversations, the next speakers continue to use honorifics, as in (1).

(1) Data Set III: As J is talking about a car accident, a waiter brings a dish to the table.

1 J(m): *asoko no atari de ushiro kara mou dakara mou*[3]

[3] Transcription conventions are as follows: ADD.HON=addressee honorific; CLF=classifier; COP=copula; GEN=genitive; HUM=humble form of referent honorific; LOC=locative; NOM=nominalizer; O=object; PST=past; QT=quotative; RES=respectful form of referent honorifics; SFP(M.M.)=moderately masculine form of SFP; SFP(M.F.)=moderately feminine form of SFP; SFP(N)=neutral form of SFP; SFP(S.F.)=strongly feminine form of SFP; SFP(S.M.)=strongly masculine form of SFP; SUB=subject; TOP=topic; VUL=vulgar form; ?=rising intonation; (( ))=nonverbal movement and scenic detail.

2 there GEN around LOC behind from well so well

3 *sakeyoi unten dayone*
4 drunken driving COP:SFP(M.M.)
5 'Around there, well, a drunken driver came from behind, you know?'

5 W(m):((bringing a dish)) *shitsurei shi**masu***
6 excuse me do:ADD.HON
7 'Excuse me.'

8 *shin jaga ni nari* ***masu***.
9 new potato LOC become **ADD.HON**
10 '(These) are roasted potatoes.'

11 C(f): *shin jaga* ***desu*** ((moving the dish to the center of the table))
12 new potato COP:**ADD.HON**
13 '(These) are roasted potatoes.'

14 J(m):*sorede tsuitotsu*
15 then collide
16 'Then, (there was) a collision.'

17 *kore tanonda no* ***watashi*** *nande ikko* ***itadaki masu***
18 this ask.PST GEN **I** thus one:CLF eat:**HUM ADD.HON**
19 'I was the one who ordered this dish, so please allow me to take one.'

20 N(f): *haai douzo* ***itadaite kudasai***
21 okay please eat:**HUM** please:**ADD.HON**
22 'Sure, please eat!'

23 J: *sonde sono kazoku ga sanzan to o uttaeteru nimokakawarazu*
24 then the family SUB a lot MetropolitanO appeal even though
25 'Then although the family repeatedly appealed to the city assembly,'

26 *kekkou karui kei ni nacchatta n dayone*
27 rather light sentence LOC end up:PST NOM COP:SFP(M.M.)
28 '(the driver) ended up receiving a rather light sentence, you know?'

Before analyzing this text in detail, I should emphasize that addressee honorifics are not expected in informal peer group conversation like this. Addressee honorifics are normally used in formal situations, such as lectures and public interviews, and in interpersonal relationships involving distance or hierarchy, such as first-time encounters and conversations with older and superior people. Because it is uncommon to use addressee honorifics in informal speech among friends who are of equal status, the use of addressee honorifics creates contextual cues and brings indexical meanings into focus.

In lines 1 and 3, J talks about a car accident in Tokyo, using a plain form of the copula with the sentence-final particle *dayone*. In lines 5 and 8, a waiter brings a dish of potatoes to the table, using the addressee honorific, *masu*. The use of addressee honorifics is always necessary in service encounters like those between waiters and their customers. In line 11, C moves the dish to the center of the table and introduces the dish with the addressee honorific *desu*. As if playing a waitress's role, she echoes the waiter's addressee honorifics. In line 14, J continues his narrative about the car accident without any honorifics.

But when J asks for permission to eat one of the potatoes in line 17, he speaks very quickly, using a humble form of referent honorific *itadaki* and the addressee honorific *masu*. Not only does he use these honorific predicates but he also uses *watashi* as the first person pronoun. This use of *watashi* is notable, because elsewhere in the data, he and another male speaker invariably use a vulgar form of the first person pronoun *ore*, common in informal male speech. But *watashi* is a polite form of the first person pronoun: a form that male speakers use in formal situations (Ide 1990). J's marked use of the first person pronoun (his choice of *watashi*) signifies a change in register from informal to formal and polite in line 17, although J is not playing a waiter's role (as C does in line 14). Then, J's use of the polite form of the pronoun *watashi* could be perceived as a strategy to minimize the risk of his face-threatening act (unlike his earlier narrative): he is asking others if he can eat one of the potatoes. Because of this inherently face-threatening act, J in line 17 has more need for politeness than he did during his earlier narratives. This could be the reason why J switches to a formal register, using the formal form of the first person pronoun and honorifics. Yet in other similar contexts in which he asks a permission to take food and eat, he does *not* use honorifics. Since honorifics do not appear in other similarly face-threatening contexts, the face-threatening nature of the utterance in line 17 is probably not the cause for his use of the formal form of the pronoun and honorifics. Given these circumstances, it is most likely that the previous two speakers' use of honorifics trigger J's use of honorifics in line 17. J echoed them with respect to the same conversational topic. The dialogicality of honorific use motivated by conversational topic becomes critical later in this segment, when the conversational topic shifts.

In line 20, N encourages J to eat, using honorifics *itadaite* and *kudasai*. Here, N makes a mistake of using referent honorifics. In talking about the addressee's action of eating, N would be expected to use the respectful form of the verb 'to eat' *meshiagatte*, not the humble form *itadaite*, although in the follow-up interview N claimed that she was unaware of her incorrect honorific usage. Instead of using honorifics (incorrectly), she could have said something more informal such as *haai tabete tabete* 'sure, eat, eat' without using honorifics. Even though the honorific form that N uses in line 20 is grammatically and pragmatically incorrect, N, at least, succeeds in repeating the same word used by J in line 17 and in attuning to the previous speakers' uses of honorifics.

In lines 23 and 26, when J switches the conversational topic back to his narrative about the car accident, his utterances no longer include honorifics. He ends his turn with *dayone*, just as in line 1. This suggests that honorifics are introduced according to conversational topic and triggered by previous instances of addressee honorifics. When the topic changes, honorifics disappear. Even if the conversational topic remains constant, the dialogic use of honorifics usually lasts for no more than four to five turns.[4] If honorifics are used beyond that limit, informal peer group conversations would sound unnatural, uncoordinated, and therefore awkward to participants.[5]

The use of referent honorifics also illustrates dialogic interaction.

(2) Data Set II: G is talking about her baby and M asks G about how happy G's parents are.

1	M:	***go-ryoushin***	*ga*	*mou*	*taihen*	*janai?*	
2		**RES**-parents	SUB	already	big deal	SFP(M.F.)	
3		'Your parents are very happy, aren't they?'					
4	G:	***haha***	*wa*	*ne*	*tomokaku*	*mou nanka*	*ne*
5		mother:**HUM**	TOP	SFP(N)	anyway	well like	SFP(N)
6		'My mother has not changed, but'					
7		*genkakudatta*	***chichi-oya***	*ga*	*konna*	*mejiri*	*sagete*
8		strict:COP:PST	father-parent:**HUM**	SUB	this	corner of eyes	lower
9		'my father who used to be very strict has become so sweet.'					

In this exchange, dialogic interaction occurs according to the conversational topic about G's parents. Once a referent honorific is used in line 1, subsequent instances of referent honorifics seem to become unmarked. After M asks G about G's parents with the subject honorific ('respectful') form of the word 'parents' *go-ryoushin*, G seems to be compelled to respond to M with the humble form of 'mother' *haha* and a relatively humble form of 'father' *chichi-oya*.[6] At first glance, their use of referent honorifics seems to be marked in this informal conversation. But once M expresses her deferential attitude to G's parents in referent honorifics, the markedness of honorific expressions changes. It would be

[4] For another example, see Takekuro (2002: 203-204).

[5] These are, however, only my own informal observations. To make a truly definitive claim, one would observe and record many more dialogic instances of honorific usage.

[6] In Standard Japanese, *chichi* is the most humble form available to refer to one's own father. Adding *-oya* 'parent' to *chichi* is less humbler than *chichi*. But *chichi-oya* is still humbler than other forms such as *oto-san/to-san* 'dad' or *papa* 'papa'.

marked for G to use a non-humble form to respond to M for making differences in their registers. In this sense, G's use of the humble forms *haha* and *chichi-oya* is somewhat unmarked. The use of honorifics is generally considered marked in informal conversations, but once an honorific is introduced into a context to activate a particular contextual aspect, the use of honorifics in the following turns can become unmarked even in informal conversations.

2.2. SFPs

Once a speaker uses a strongly gendered form of SFP, other speakers also dialogically resonate with analogous strongly gendered SFPs.

(3) Data Set I: The speakers are complaining about their professors.

1	T:	*omae*	*fuzaken*	***nayo***	*tte*	*omotta*	*yo*
2		you:VUL	kidding	**SFP(S.M.)**	QT	think.PST	SFP(M.M.)
3		'I thought you shouldn't be so ridiculous.'					
4	K:	*omou*	***yona***				
5		think	**SFP(S.M.)**				
6		'I think so.'					
7	I:	*kocchi*	*mo*	*ganbatta*	*n*	*da**zo***	*ttsu*	*no*
8		this	too	do the best:PST	NOM	COP:**SFP(S.M.)**	QT:say	NOM
9		'I was like "we were doing our best!"'						

In (3), when T uses the strongly masculine form of SFPs *nayo* to express her annoyance at her professors, K and I then also use strongly masculine forms. Although these forms are not statistically common for young female speakers to use even in informal conversations, T's use triggers more forms in the same gendered category from K and I. The same happens with strongly feminine forms of SFPs, as in (4) where speakers use them one after another.

(4) Data Set II: Remarking on the changes in the content of their conversation.

1	K:	*nanka*	*atashitachi*	*mo*	*wadai*	*ga*	*otona*	*ni*	*natta*	*yone*
2		well	we	also	topic	SUB	adult	to	became	SFP(N)
3		'Well, the topics of our conversation have become grown up.'								
4		*sou*	*ima*	*shimijimi*	*omocchatta*	*yo*				
5		yes	now	really	think.PST	SFP(M.M.)				
6		'Yes, I just really felt so.'								
7	G:	*soryaa*	*gonen*	*mae*	*to*	*wa*	*chigau*	***wayo***		
8		of course	five years	ago	QT	TOP	different	**SFP(S.F.)**		
9		'Of course, we don't talk like we did 5 years ago.'								

10 M: *minna* *otona* *ni* *natta* ***noyo***
11 everyone mature to become.PST **SFP(S.F.)**
12 'Everyone became mature.'

13 T: *soshite* *haha* *ni* *natte* *iku* ***none***
14 then mother to become go **SFP(S.F)**
15 'Then we become mothers.'

While each instance takes a different linguistic form in both (3) and (4), they are still in the same gendered categories. Using the form in the same category, which is initially considered marked due to the statistically infrequent usage, thereafter becomes unmarked in the interactive context, because the previous instance sets a new norm of interaction for following instances.

3. Discussion

3.1. Dynamic changes of indexicals

My examples of dialogic interaction demonstrate the use of linguistic forms as indexical signs. The use of linguistic forms such as honorifics and strongly gendered forms of SFPs can be highly marked in informal peer group conversations. When speakers engaged in informal conversations use marked linguistic forms, they convey additional information about the speaker's perception of the interactive context. For instance, honorifics give a sense of formality, politeness, or irony to the interaction, while strongly masculine or feminine forms of SFPs convey coarseness or gentleness, respectively. However, from the perspective of dialogic interaction, the first instance of marked linguistic forms sets a new norm of interaction and brings the next instance of marked linguistic forms into conformity with the new norm in the situated interactive context. Subsequent instances no longer seem anomalous, rather unmarked. On the other hand, if subsequent speakers do not use marked linguistic forms that are introduced in prior turns, the use of unmarked linguistic forms could be perceived as a violation of the new norms in the situated interactive contexts, conveying more information than the use of marked linguistic forms.

Initially, the choice of honorifics and SFPs is governed by socially motivated norms of interaction that are relatively "fixed": the formality of the situation, participants' status hierarchy, and the speaker's gender. But whether each instance of usage is interpreted as unmarked or marked depends on its immediate interactive context, such as the conversational topic and participants' role structure. In other words, what speakers experience as "marked" or "unmarked" utterances may shift based on prior context. After one speaker's usage foregrounds other speakers' options, the others discern a change in indexicals and make linguistic alignment by their own usage of analogous linguistic forms for the next four or five turns. The choice of linguistic forms is on the one hand defined by the immediate interactive context, but on the other hand the interactive context is also defined by the choice of linguistic forms (cf. Duranti 1992).

During the processes of interaction and interpretation, the distinction between unmarked and marked aspects of indexicals is not static, but dynamic and constantly negotiated in situated contexts. Even though interaction is largely based on socially motivated norms of language usage, the norms are nonetheless fluid. In fact, if norms of interaction are strictly fixed, it is difficult to explain speakers' creativity in using language to alter interpersonal relationships and achieve special effects such as irony, sarcasm, or humor (Agha 1994: 288). The fluidity of norms of interaction allows meanings of linguistic forms and social relationships to be creatively negotiated in the interactive contexts like individual conversations. In saying this, however, I do not mean to underestimate the existence of socially motivated norms. The fact that the dialogic use of honorifics lasts for only four or five turns indicates that speakers are aware of "markedness" of using honorifics in informal conversations with respect to socially motivated norms. Thus, indexical meanings of linguistic forms arise out of the dynamic interplay between broader social norms of interaction and the situated interactive context.

3.2. Dialogic interaction in light of Goffman's face-work

In providing these socio-cultural accounts of dialogic interaction, I draw on Goffman's (1956) notion of "deference" and "demeanor" as conceptual frameworks. Deference and demeanor both involve respect, for the other and for oneself. Deference is defined as what is owed to the other and demeanor as what one owes to oneself. Using these notions, in this section, I discuss the socially significant meanings of dialogic interaction.

In informal conversations, speakers have a choice between dialogic and non-dialogic interaction. Although speakers always have the choice of using less marked forms, they "may" choose to make use of previous speakers' marked linguistic forms. Particularly when the degree of markedness is high, dialogicality through the use of analogous linguistic forms or register maintains coherence in interaction. Dialogic interaction demonstrates a speaker's acknowledgement that participants are interacting on common ground. In this way, speakers show deference to one another, because dialogic interaction increases feelings of connectedness among participants. Additionally, dialogic interaction conveys metamessages about the speakers themselves, who display their demeanor as socially engaged members of the group. Dialogic interaction helps parties to establish common ground; deference enhances parties' regard for one another's self-respect; and demeanor creates parties' regard for their own. Failure in any of these respects might be experienced by participants as 'having nothing in common' (Hymes 1986: 82). Therefore I interpret dialogic interaction as a way of ensuring that one is engaged properly in discourse and allowing all speakers to signal their identity and membership in the group.

4. Conclusion

As an analysis of the use of honorifics and SFPs in dialogic interaction, this study suggests that these linguistic forms are multifunctional and multiply indexical. Even in highly marked uses, linguistic forms may have the status of givens, and therefore become unmarked. As their functional and pragmatic dimensions of usage are subject to negotiations between socially motivated norms of interaction and the immediate interactive context, we cannot take indexical meanings as pre-given or predictable a priori. This study, however, is just a beginning. More studies on the indexical properties of natural language are necessary if we are to come to understand the dynamic interplay between linguistic forms and interactive context.

References

Agha, Asif. 1994. Honorification. *Annual Review of Anthropology* 23: 277-302.

Bakhtin, Mikhail. 1981[1935]. In M. Holquist (ed.) *The dialogic imagination*. Austin: University of Texas Press.

Du Bois, John. 2003. Self other boundary: Stance & consequence in the construction of intersubjectivity. Presented at Kroeber Anthropological Society at UC Berkeley.

Duranti, Alessandro. 1992. Language in context and language as context: The Samoan respect vocabulary. In A. Duranti & C. Goodwin (eds.) *Rethinking context: Language as an interactive phenomenon*, 77-99. Cambridge: Cambridge University Press.

Goffman, Erving. 1956. The nature of deference and demeanor. *American Anthropologist* 58: 473-502.

Goffman, Erving. 1974. *Frame analysis: An essay on the organization of experience*. New York: Harper and Row.

Gumperz, John. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Hymes, Dell. 1986. Discourse: Scope without depth. *International Journal of the Sociology of Language* 57: 49-89.

Ide, Sachiko. 1990. How and why do women speak more politely in Japanese? In S. Ide & N.H. McGloin (eds.) *Aspects of Japanese women's language*, 63-79. Tokyo: Kuroshio Publishers.

Jakobson, Roman. 1960. Concluding statement: Linguistics and poetics. In T. Sebeok (ed.) *Style in language*, 350-377. Cambridge: MIT Press.

Jakobson, Roman. 1966. Grammatical parallelism and its Russian facet. *Language* 42: 398-429.

Lyons, John. 1977. *Semantics 1*. Cambridge: Cambridge University Press.

Morris, Charles. 1938. Foundations of the theory of signs. In O. Neurath, R. Carnap & C. Morris (eds.) *International encyclopedia of unified science*, 77-138. Chicago: University of Chicago Press.

Okamoto, Shigeko. 1997. Social context, linguistic ideology, and indexical expressions in Japanese. *Journal of Pragmatics* 28: 795-817.

Okamoto, Shigeko & Shie Sato. 1992. Less feminine speech among young Japanese females. In K. Hall, M. Bucholtz & B. Moonwomon (eds.) *Locating power: Proceedings of the second Berkeley Women and Language Conference*, 478-488.

Peirce, Charles. 1955. Logic as semiotics: The theory of signs. In J. Buchler (ed.) *Philosophical writings of Peirce*, 98-119. New York: Dover Publications.

Silverstein, Michael. 1976. Shifters, linguistic categories, and cultural description. In K. Basso & H. Selby (eds.) *Meaning in Anthropology*, 11-55. Albuquerque: University of New Mexico Press.

Sturtz, Cindi. 2001. *Danseigo da zo*! Japanese men's language: Stereotypes, realities, and ideologies. Unpublished Ph.D. dissertation, UC Davis.

Takekuro, Makiko. 2002. Indexicality and socialization: Age-graded changes of young Japanese women's speech. In K. Kataoka & S. Ide (eds.) *Culture, interaction and language*, 195- 214. Tokyo: Hitsuji Publishers.

Makiko Takekuro
1203 Dwinelle Hall
Department of Linguistics
University of California, Berkeley
Berkeley, CA 94720-2650

makikot@socrates.berkeley.edu

The causatives in Sun Hongkai's Anong: language death and rapid restructuring

GRAHAM THURGOOD
California State University Chico

Introduction. The Anong people, numbering approximately 6500, mainly live in Fugong County in the Nujiang Lisu Autonomous Prefecture of Yunnan Province (roughly 98.7°E and 27.1°N) on the tablelands on both sides of the Nujiang River. To the west is Myanmar, where the closely-related Rawang live; to the north is the Tibetan Autonomous Region.

Anong, an autonym, is a Tibeto-Burman language closely-related both to Trung [tə̆rùŋ] (Chinese exonym: Dulong; population c. 6000 (LaPolla 2000:282-283)) and to Rawang, with the details of higher-level relationships remaining to be sorted out. The Anong people live surrounded by Lisu, Chinese, and Bai speakers. In fact, the majority of the Anong people no longer speak Anong, most having shifted either to Lisu or Chinese, with a few having shifted to Bai. Ethnic Anong who still speak Anong are now found primarily in the villages, such as Mugujia, Kashi, Muleng, Lagagong, Aniqia, and Lahaigong, which belong to the Mugujia Township of Shangpa Town in Fugong County.

Causatives. Anong causatives are one of many linguistic subsystems that have been totally restructured in the last forty years. The data that Sun Hongkai recorded forty years ago retains the older, largely inherited, and quite familiar Tibeto-Burman system, which marks most causative forms by prefixing *sɯ*^{55}- (or *ɕi*^{31}- before palatals) to the basic root. The newer contemporary system, recorded in 1999, is much messier: the older, easily-segmentable prefix + root combinations have undergone considerable assimilation, with the former prefixes fusing with the initial of the root; this has resulted in a system in which the simplex forms (i.e., non-causative forms) are distinguished from the causative forms by differences in the root-initial consonants.

While the basic outlines of these phonetic changes have been known to Tibeto-Burmanists for some time (e.g., Matisoff 1972; Thurgood 1977, 1981; and so on), some of the phonetic details are nonetheless of interest. In a recent paper, Sun (1999a:355-356) notes that in 1983, the Anong used by an older speaker (70 years old) marked causatives with both the inherited Proto-Tibeto-Burman (PTB) prefix and with transparently-related changes in the verb roots—thus, the older system had been augmented by subsequent assimilation. However, doing additional fieldwork in 1999, just sixteen years later, Sun's reports (1999a:355-356) that the causative system used by his new language assistant did not include many of the forms used by the 1983 language assistant. He states that, while some speakers claimed to recognize the older forms, others did not; some suggested that the system had undergone simplification. Table 1 exemplifies one of the differences between the oldest, most conservative speakers and others.

Table 1 shows the unprefixed forms in the first column; the second column shows the causative forms used by the older speakers, and the third column shows the causative forms of the middle-aged speakers. The phonetic correspondences are straightforward: in the causatives, the older voiced obstruents have become voiceless aspirated obstruents and the older nasals (sonorants) have become voiceless.

Table 1: Anong causatives (older vs. middle-aged speakers)

verb	causatives (older speakers)	causatives (middle-aged speakers)	gloss
dʑɯŋ^{55}	ɕi^{31}dʑɯŋ^{55}	ɕtɕɯŋ^{55} > tɕʰɯŋ^{55}	be broken
ȵi^{55}	ɕi^{31}ȵi^{55}	ȵ̊i^{55}	know
ga^{55}	sɯ^{55}ga^{55}	kʰa^{55}	be broken
ba^{55}a^{31}	sɯ^{31}ba^{55}a^{31}	pʰa^{55}a^{31}	be white

Source: Sun (1999b:195); transcription slightly modified.

The phonetics of the form *ɕtɕɯŋ^{55}* 'be broken', with its prefix fused to the onset of the root are, to my knowledge, unattested elsewhere, although hardly unexpected. In Anong *ɕtɕɯŋ^{55}*-like forms have developed into voiceless aspirated onsets; in other Tibeto-Burman languages they have instead sometimes become voiceless unaspirated. Still none of this is particularly unexpected even if the Anong phonetic record is somewhat richer than is typical for this sort of change.

The developments themselves are fairly natural changes: the prefixal *sɯ^{31}-* losing its unstressed vowel, fusing with the following root, and then devoicing a root initial stop or nasal before dropping is widely attested. The fusion of the prefixal *ɕɯ^{31}-* replacing the root initial lateral is not all that uncommon, nor is the complementary relationship between the prefixed unaspirated voicieless root initial and the subsequent aspirated voiceless root initial once the prefix has dropped. In fact, the changes themselves are so natural, it is not the changes that took place, but rather that the system was stable for so long. Thus, perhaps it is not the changes that require explanation, but instead it is the long-term stability.

Two things set the Anong data apart: a precise time frame and clear, richly-attested data. In Sun's 1960 Anong descriptions the older system is largely intact, but in his recent 1999 work, a new, fully-restructured system has taken its place. Further, the data recorded between these two dates show intermediate stages, with the stages thus requiring a minimum of unattested extrapolation. Thus, the major unresolved questions revolve neither around the phonetic details nor the time frame, but focus on activation, that is, Why now? After all, the 1960 system is essentially the same system that dates back as far as Proto-Tibeto-Burman—at the very least some two thousand years and more than likely much older. What are the changes in the last forty years of Anong history that provided the impetus for restructuring?

The answers correlate with recent changes in the social setting. While the earlier structure of Anong of course places some constraints on the restructuring of the Anong causative, the primary are in the social setting. The most obvious indication of this comes from the comparison of Anong causatives with the causatives of closely-related Trung: Table 2 shows that the Trung (Dulong) causatives are strikingly similar to the older causatives attested in older Anong speakers, while the newer causatives used by middle-aged (and younger) speakers are now quite distinct.

Table 2: Trung (Dulong) causatives vs. Anong causatives

Trung (Dulong)		Anong			
verb	causative	verb	older causative	newer causative	
lɯp^{55}	tɯ^{31}lɯp^{55}	lim^{55}	ɕi^{31}lim^{55}	ɕim^{55}	bury
dɯm^{55}	sɯ^{55}dɯm^{55}	dim^{55}	ɕi^{31}dim^{55}	ctim^{55} > tʰim^{55}	collapse
ŋɯ^{53}	sɯ^{31}ŋɯ^{53}	ŋɯ^{55}	sɯ^{31}ŋɯ^{55}	ŋ̊ɯ^{55}	cry

Source: Sun (1999b:194-195)

In short, although both languages were extremely similar just 40 years ago, while Trung has undergone only moderate change since then, Anong has completely restructured.

As for the causatives themselves, notice that the older Anong causatives, like the causatives of Trung, mark the causative by the prefix *sɯ^{55}-* (or *ɕi^{31}-* before palatals). The newer causatives, unlike in Trung, are essentially lexical—for those without a background in historical linguistics, the causative forms need to be learned one at a time, as in Lisu, a language most Anong speakers are more fluent in than they are in Anong. Note, however, the similarity to Lisu only tells us a small part of the why the restructuring occurred as it did: first, Anong speakers fluent in Lisu had managed to keep both systems separate for quite some time before the restructuring of the Anong causatives began, and, second, the most that could have been adapted from Lisu is the principle of lexical suppletion for causatives—the forms themselves, of course, were not borrowed.

Speakers, numbers, and fluency. The apparent similarity between the population figures for Anong (c. 6500) and those for Trung (c. 6000) is deceptive: they reflect ethnic affiliation more than the number of speakers. In fact, it is here where one of the more striking differences between the Anong and Trung can be seen. Trung is still a viable language, used for a range of purposes, and with younger speakers. Anong is a language on the edge of extinction, under intense pressure from Lisu as a local language and from Chinese as a language of wider communication, with only 400 or so remaining Anong speakers, and of these, only roughly 12% are fluent.

In a 1999 paper, Sun (1999a:353-355) reports his survey of some 104 people from 25 households—roughly a quarter of all the Anong speakers, located in Mugujia village, the area where Anong is the most concentrated. The survey checked their fluency in Anong, in Lisu, and in Chinese, in terms of four levels of linguistic competence, here called Levels A, B, C, and D. These levels are described below, as they apply to Anong.

Level A—Fluent: Level A speakers are able to engage in everyday conversation as well as describe food production and daily living and are able to use over 3,000 basic items of vocabulary.

Level B—Semi-fluent: Level B speakers can engage in everyday conversation, but not very fluently and able to use around 1,000 basic words. Level B speakers are better in their second language than in Anong.

Level C—Limited: Level C speakers are only able to use daily greetings or a few everyday phrases and, thus, are unable to use Anong to express themselves adequately and completely. Their Anong frequently includes words from their second language. They only use a limited number of Anong vocabulary items, with much of the basic vocabulary missing. Their pronunciation of Anong is inaccurate. In contrast, they are very fluent in their second language.

Level D—Non-speakers: Level D respondents have essentially lost their mother tongue, although some can still understand a little, but they can no longer speak their mother tongue. In fact, some have completely lost their Anong.

Table 3 puts Sun's (1999a:353-355) numbers together in table form. Examination of the Anong column shows some 62 individuals, roughly 60%, are still fluent in Anong (using the criteria above). However, Sun (1999a:354) writes, outside of Mugujia village, there are virtually no fluent Anong speakers left! Essentially, these 62 are the last fluent speakers. And, Sun found some 19 semi-fluent speakers in the Mugujia sample. Sun (1999a:254) notes that, although there are no fluent speakers left, semi-fluent speakers are still found outside of the Mugujia area, suggesting that there are roughly 280 of these outside of Mugujia. In sum, there were as of 1999 roughly 62 fluent Anong speakers and perhaps a little over three hundred semi-fluent speakers remaining.

Table 3: Fluency rates in Anong, Lisu, and Chinese.

	Anong		Lisu		Chinese	
Proficiency	#	%	#	%	#	%
A. Fluent	62	59.6%	96	92.3%	13	12.5%
B. Semi-fluent	19	18.2%	8	7.7%	17	16.3%
C. Limited	14	13.4%	0	0.0%	19	18.3%
D. Non-speakers	7	8.8%	0	0.0%	55	52.6%

However, as Sun (1999a:353-355) observes, a closer look at the 62 fluent speakers of the 104 Anong speakers in Mugujia village shows how close the language is to being lost. As the next column indicates, of the 104 Anong surveyed (including the 62 fluent speakers), 96 of them—all but 8—are fluent in Lisu. All 8 of those not fluent in Lisu were elderly people who rarely went out. Thus, as Sun notes, that even in the village where the highest concentration of Anong speakers was found, the proficiency and level in Lisu has already surpassed that of Anong. Further, as he notes, almost all the Anong, including most of the fluent speakers, can speak Lisu and most can speak it better than Anong.

The situation with Chinese, the final column, seems quite different. While Sun's numbers show some 30% of the Anong as fluent or semi-fluent in Chinese, it is typically a local dialect different from the Kunming dialect rather than Putonghua. Those who were fluent in Chinese were almost all middle-aged or younger and had lower secondary school level educations or higher and for the most part those who were semi-fluent frequently had either a primary school education or had at least attended school. As a consequence of this schooling, fluent and semi-fluent speakers of Chinese had more opportunities for interaction outside the village.

Usage patterns. The usage patterns correlate with the loss of frequency. Sun (1999a:354) writes that Mugujia village has both the highest concentration of Anong speakers and, not surprisingly, the highest usage of Anong. As Sun observed, in Mugujia village it was possible to hear conversations in Anong, although even here most people spoke in Lisu.

Outside of Mugujia village Anong usage falls off sharply. Although Anong speakers also live in other villages around Mugujia—Kashi, Muleng, Lagagong, Aniqia, and Lahaigon, in these areas the Anong often are less concentrated, living interspersed with concentrations of Lisu or with people of other nationalities. In these areas, Sun (1999a:354) notes, there are essentially no remaining Anong speakers who are fully fluent. In these areas, Sun (1999a:354) counted some 280 semi-fluent speakers, with most being elderly people in their 60s or 70s. Of the slightly younger speakers, Anong in their 50s or 60s, only a handful could still speak their mother tongue even at the semi-fluent level. Sun (1999a:354) attributes this to the lack of opportunity to use Anong in their daily lives; most, he notes, only used a few sentences and then only use these when speaking to the elderly. In these areas, Lisu was the language used for common communication.

Demographics, usage, and restricted access. As Sun suggests (1999a:354), not only is Anong with its relatively complicated phonological and grammatical systems difficult to learn, but the patterns of interaction severely restrict usage. The Anong have interacted with various ethnic groups in the Nujiang River area, but especially with the Lisu. In conversations, unless all individuals spoke Anong fluently, the common language would, of course, have been Lisu. In the case of intermarriage where an Anong had married someone from another group, the same pattern held: even though only one member of the family did not speak Anong, the common language was Lisu, with the non-Anong only very rarely learning some Anong. Sun (1999a:354) observed that even in the rare villages where the Anong were the majority, few non-Anong learned Anong.

Sun (1999a:354) illustrates the typical language in the case of intermarriage by describing the patterns of his Anong language assistant. His assistant was fluent in Anong, but his wife was Lisu. They had been married 30 years, but even after 30 years, she spoke only a little Anong and even that rarely. Lisu was the common language in the family: the three daughters and the son, like the mother, only understood a little Anong, and none could speak it. Occasionally, the father would use a little Anong at home, but they would still answer in Lisu.

In sum, Anong tends to be restricted to situations where all participants speak it fluently: the default language is virtually always going to be Lisu. Thus, only fluent Anong speakers tend to have access to Anong; those less fluent in Anong only have very restricted access to the language.

Conclusions and speculation. One mechanism involved in the loss of Anong is simple replacement, generally by Lisu, a development undoubtedly exasperated by limited usage and limited access to Anong . Anong forms that are used infrequently are more easily forgotten; Lisu forms that are used frequently are more readily retrieved. All Anong speakers, even the most fluent, probably show some evidence of simple replacement; Sun cites the example of one of his best language assistants substituting the Anong form for ‘very’ with its Lisu equivalent. With less frequent usage, replacement should go up; with less proficiency, replacement should go up.

Other more minor factors also come into play. Certainly, although more marginal, one factor is attitudinal. Sun (1999a) notes that many of the best educated Anong are relatively indifferent to the impending loss of Anong, expressing the view that not only is this the general trend but also noting that they realize that there is little they could do about it in any case. It is worth noting that, while Anong is their first language, the language they are losing is neither their only language, nor their most useful language, nor even the language that they are most fluent in; almost all Anong speak Lisu better than Anong and use it for more purposes.

The shift to Lisu was well underway before the founding of the People's Republic of China in 1949; in fact, it must have been well underway at the time of Barnard's 1934 (1934:89) grammatical sketch and glossary of closely-related Rawang which lists as the Rawang term for Lisu the word Anung. However, it is only in the last 30 years or so that the decline in the numbers of speakers and the increasingly restrictive usage patterns have produced the massive restructuring of Anong, not just in its causatives, but in all its systems.

Thus, it is obvious that the changes are a response to language contact and subsequent changes in usage patterns, however, to say this is to point out a correlation rather than to explain anything. Nor is it much of an analysis to invoke 'assimilation to Lisu' as an explanation. Where Lisu differentiates causative and non-causatives, it is with suppletion. If so, what would Anong speakers be said to be assimilating from Lisu? Certainly not the specific words. Certainly not the process.

It is, however, true that Lisu is exerting an influence on the phonology of Anong, but while Lisu influence might account the directionality of the phonological changes, it does little to account for the pace and timing of the restructuring. The bulk of the explanation for the rapidity of the changes lies outside of the phonology of Lisu.

It is clear that the older generation of Anong speakers is the last generation to successfuly learn Anong. The middle-aged and younger speakers have not done so. And, as already observed, this failure to learn Anong correlates with the rapid restructuring and with the increasingly restricted usage and access. In this sense, access to Anong has become too restricted to be successfully passed on; the middle-aged and the younger Anong do not seem to be learning it.

None of this seems controversial. However, none of this accounts —directly—for the rapid restructuring either. The key to understanding lies in the nature and direction of the restructuring. Our examination of the Anong causative restructuring reveal a series of changes that, although they occurred unusually rapidly, were quite natural. For instance, all of the following developments are fairly natural changes: a prefixal *sɯ*^{31}- losing its unstressed vowel, fusing with the following root, devoicing a root initial stop or nasal, and then dropping.

In fact, the changes look to natural changes—the only thing requiring an explanation is the rapidity with which they took place. Thus, returning to a suggestion made earlier in this paper, these changes are so natural, it is necessary to explain not only why the changes have not occurred elsewhere, for instance, in the strikingly-similar Trung data, but also why they haven't occurred as rapidly.

The answer is that normal transmission undoes both phonetic and morphological changes in children and, at least, in part in adult learners. Young and old learners regularly 'repair' their own phonetically-motivated changes to make them conform to what they perceive as more desirable pronunciations. The repairs in morphology are a little more sophisticated but produce a similar result. Initially, the learner seems to simply acquire morphology as a series of individual tokens. At a later stage, learners often recognize a morphological pattern, and, in part restructure earlier forms on the basis of their generalizations—in the case of causatives by putting the prefix before the basic root. Later, some phonetically-motivated changes that occurred quite naturally in the token stage of acquisition are at least partially undone when the learner recognizes the morphological pattern, in the case of causatives putting the prefix before the basic root—thus, connecting in some sense the unprefixed root with the prefixed root. This second stage, generalizing the pattern, has a tendency to undo the phonetic effect of the prefix on the root initial: the causative is reanalyzed each generation as a prefix plus the root. In Anong, however, the youngerr learners of Anong seem to have never gotten beyond the stage of learning individual tokens—essentially because they lacked sufficient access to language. The result was the connection between the basic root and the causative root was lost—along with the constraint that that connection imposed on phonetic and morphological restructuring. The fact that Lisu uses a lexical approach too probably played a minor role, but the major impetus to restructuring came from increasingly restricted access to the language—a restriction that made the morphological patterns too rare and thus too obscure to learn. The phonetic tendencies explain how the changes took place; the loss of sufficient access to the language explains why the changes occurred in the last 40 years.

Acknowledgements

The research in this paper is supported in part by National Science Foundation Grant No. 60232 Endangered Languages in China. The Anong data is from Sun Hongkai's work, a participant in the grant. The data is part of a remarkable collection study of Anong, a study spread over some 40 years, full of treasures. I wish to thank Fengxiang Li for his invaluable insights and suggestions. I also wish to thank a number of conference participants for feedback, some of which is still being followed up (David Mortensen, Larry Hyman, Johanna Nichols, and others).

References

Barnard, J.T.O. 1934. *A handbook of the Rāwang dialect of the Nung language, containing a grammar of the language, colloquial exercises and a vocabulary with an appendix of Nung manners and customs.* Rangoon: Superintendent, Government Printing and Stationery, Burma.

LaPolla, Randy J. 2000. 'Valency-changing derivations in Dulong/Rawang.' *Changing Valency: Case Studies in Transitivity,* ed. by R. M. W. Dixon & Alexandra Y. Aikhenvald, 282-311. Cambridge: Cambridge University Press.

LaPolla, Randy J. 2003. Dulong. In Graham Thurgood and Randy LaPolla 2003, pp. 674-682.

Matisoff, James A. 1972. *The Loloish Tonal Split Revisited.* Research Monograph No. 7, Center for South and Southeast Asia Studies, University of California, Berkeley. 88 pp.

Sun Hongkai. 1988. Notes on Anong: a new language. Translated by Fengxiang Li. *Linguistics of the Tibeto-Burman Area* 11:27-63.

Sun Hongkai. 1999a. Notes on Anong: Observations of its path of gradual decline and loss]. *Zhongguo Yuwen* 272.5:352-357.

Sun Hongkai. 1999b. The category of causative verbs in Tibeto-Burman languages. *Linguistics of the Tibeto-Burman Area* 22.1:183-199.

Thurgood, Graham. 1977. Lisu and proto-Lolo-Burmese. *Acta Orientalia* 38:147-207.

Thurgood, Graham. 1981. *Notes on the origins of Burmese creaky tone.* Monumenta Serindica Monograph Series No. 9. Institute for the Study of Languages and Cultures of Asia and Africa. Tokyo. 94pp

Thurgood, Graham and Randy J. LaPolla (eds.). 2003. *The Sino-Tibetan languages.* Routledge Language Family Series. London and New York: Routledge.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann and Yakov Malkiel (eds.). *Directions for historical linguistics.* Austin and London: University of Texas Press. Pp. 95-188.

Zhangmianyu Yuyin he Cihui (= ZMYYC) Tibeto-Burman phonology and lexicon. 1991. China Social Sciences Press.

Graham Thurgood
English Department #830
California State University, Chico
Chico, California 95929-0830

gthurgood@csuchico.edu

Junctural and parasitic voicing in Burmese[1]

KENNETH VANBIK
University of California, Berkeley

In this paper I address some of the complications that arise concerning consonant voicing in Burmese. Burmese is typically cited in phonological work for its contrast between voiced and voiceless sonorants, especially nasals (Ladefoged & Maddieson 1994:111). The full consonant system is indicated in (1) (cf. Roop 1972).

(1) Consonants (all found in onsets of syllables)

p		t	c	k	?
ph		th	ch	kh	
(b)		(d)	(ɟ)	(g)	
	θ	s			h
		s^h			
	(ð)	(z)			
m		n	ɲ	ŋ	
hm		hn	hɲ	hŋ	
w		l	y		
		hl	hy		

Note that I have indicated voiced obstruents in parentheses, since they are only marginally phonemic in Burmese, mostly found in loanwords. On the other hand,

[1] I would like to thank Jim Matisoff, Larry Hyman, and Sharon Inkelas for their guidance and suggestions for this paper. The collection of data for this paper began in 1999 for Prof. Hyman's course: Ling.111 'Introduction to Phonology'. I would like to thank Ko Moe Zaw and Ko Maung Maung Nyo for allowing me to check with them the pronunciation of the data used in this paper. Ko Moe Zaw is a Burmese from Rangoon (Lower Burma Burmese), and Ko Maung Maung Nyo is a visiting scholar in Journalism here at UCB. He is from Mon Yua (Upper Burma Burmese). Because they were very consistent with the voicing phenomena described in the *Myanmar-English Dictionary* (MED 94), later collections were mainly done from the MED. The data collection is quite extensive, but not exhausive enough to give an accurate statistical evaluation. Needless to say, all the mistakes are mine.

derived voiced stops will be the focus of the present study. In particular, we shall be interested in the voicing counterparts indicated in (2).

(2)

voiceable	plain	p	θ	t	s	c	k
	aspirated	ph		th	s^h	ch	kh
voiced counterpart		b	ð	d	z	j	g

Specifically we will analyze the conditions under which so-called "voiceable" consonants become their voiced counterpart. (Curiously, voiceless sonorants do not undergo voicing). To understand this, we need to consider the vowel system and two types of rhymes in Burmese. As seen in (3), vowels in open syllables can be oral or nasalized (marked as *ñ*)[2], and the only type of closed syllables ends in a glottal stop (cf. Matisoff 1976:v)[3].

(3) Rhymes

Checked		Open			
i?	u?	iñ	uñ	i	u
ei?	ou?	eiñ	ouñ	ei	ou
				wei	
e?				e	o
we?				we	
ai?	au?	aiñ	auñ		
a?		añ		a	
				wa	

The first source of voiced obstruents in Burmese is schematized in (4).

(4) Junctural voicing

[voiceless obstruents] → [+voice] / V + ___ V

As seen in (4), voiceless obstruents become voiced intervocalically by a process which I term "junctural voicing". This occurs obligatorily in several grammatical contexts. First, voicing occurs on the second element of a compound (cf. 5).

[2] Following Okell 1969:6, nasalized vowels are marked by a symbol *ñ* so that tones can be marked on the vowels. Tones are marked as follows: heavy tone by grave accent [`], level as unmarked, and creaky tone by acute accent [´].

[3] Glottal stop is often considered to be a tonal feature because all syllables ending with glottal stop are pronounced with the same short high checked tone.

(5) Compounding

a.	p > b					
	pè	+	pouʔ	---->	pè-**bouʔ**	'fermented soybean'
	'peanut'		'rot/putrid'			
b.	t > d					
	khè	+	tañ	---->	khè-**dañ**	'pencil'
	'lead'		'stick'			
c.	c > j					
	pè	+	cì	---->	pè-**jì**	'lablab bean'
	'peanut'		'big'			
d.	k > g					
	naiñ	+	kweʔ	---->	naiñ-**gweʔ**	'trump card'
	'win/can'		'spot'			
e.	θ > ð					
	pàñ	+	θì	---->	pàñ-**ðì**	'apple'
	'flower'		'fruit'			
f.	s > z					
	kouñ	+	sèiñ	---->	kouñ-**zèiñ**	'raw materials'
	'goods'		'be green'			

The examples in (6) show the same phenomenon occurring in reduplication.

(6) Reduplication

a.	pyauñ	+	pyauñ	---->	pyauñ-**byauñ**	'bluntly'
	'be plain'		'be plain'			
b.	té	+	té	---->	té-**dé**	'straight-forward'
	'be straight'		'be straight'			
c.	ci	+	ci	---->	ci-**ji**	'cheerfully'
	'joyous'		'joyous'			
d.	kàuñ	+	kàuñ	---->	kàuñ-**gàuñ**	'properly'
	'be good'		'be good'			

e.	θouñ 'blow'	+	θouñ 'blow'	---->	θouñ-ðouñ	'(blow) briskly'
f.	sò 'early'	+	sò 'early'	---->	sò-**z**ò	'be ahead of'

Finally, in (7), we see that certain grammatical markers also become voiced.

(7) Grammatical markers (subject, object, politeness etc.)[4]

a. ka 'Subject (SUB) marker' MS : 257 (4)
thu. hsi **ga.** ya de
3P.G place SUB get RLS
'I got (it) from him'

b. kou 'Object (OBJ) Marker' MS: 99 (9)
kou=ei: **gou** thu **ga**. tha' te
HON=name OBJ 3P SUB kill RLS
'He killed Ko Aye'

c. pa 'Politeness (POL) Marker' MS: 126 (10)
sa. ja **ba** zou.
start PLU POL HORT
'Let's start, shall we ?'

As indicated in the informal rule in (4), the affected consonant must be preceded by a vowel. As seen in (8), voicing will not accompany the compounding process if the first word ends in glottal stop.

(8) Compounding

a.	θiʔ 'tree'	+	pouʔ 'rot/putrid'	---->	θiʔ-pouʔ	'malodorous wood-tree'
b.	cauʔ 'stone'	+	tañ 'stick'	---->	cauʔ-tañ	'slate pencil'

[4] Grammatical markers are collected from Myint Soe's dissertation (MS). Following the *Myanmar-English Dictionary* (1993), Myint Soe (cf. MS:13) explains that he represents glottal stop as a comma at the end of the syllable, and voicing is indicated by underlining the consonant. In this paper, those underlined consonants are transcribed as voiced consonants. Myint Soe also transcribes s^h as *hs* .

c.	cauʔ 'small pox'	+	cì 'big'	---->	cauʔ-cì	'kind of smallpox'
d.	pyeʔ 'miss'	+	kweʔ 'spot'	---->	pyeʔ-kweʔ	'absence'
e.	θiʔ 'tree'	+	θì 'fruit'	---->	θiʔ-θì	'fruit'
f.	cauʔ 'stone'	+	sèiñ 'be green'	---->	cauʔ-seiñ	'jade'

Similarly, as seen in (9), voicing will not occur in reduplication if the base ends in a glottal stop.

(9) Reduplication

a.	kauʔ 'curl'	+	kauʔ 'curl'	---->	kauʔ-kauʔ	'curly'
b.	cauʔ 'fear'	+	cauʔ 'fear'	---->	cauʔ-cauʔ	'in fear'
c.	souʔ 'suck'	+	souʔ 'suck'	---->	souʔ-souʔ	'(sink) without a trace'

Finally, in (10), we see that grammatical markers will not get voiced if they are preceded by glottal stop.

(10) Grammatical marker

a. ká 'Subject (SUB) marker' MS: 77 (9)

thu. hcei=htau' **ka.** nga. gaun: gou hti. nei de
3P.G leg SUB 1P.G head OBJ touch stay RLS
'his leg is touching my head'

b. kou 'Object (OBJ) Marker' MS: 264 (35)

sa=ywe' **kou** hkau' pa
paper OBJ fold POL
'Please fold the paper'

c.	pa	'Politeness (POL) Marker'	MS: 126 (10)

pyan lai' **pa** oun: me
return away POL AUX IRR
'I am leaving (going home)'

As seen in (11), loanwords also follow the above general rules of junctural voicing.

(11) Loan words

a.	kouñ	+	kà	---->	kouñ-ga	'lorry car'
	'goods'		'car'			
	[native BS]		[< ENG]			
b.	s^{h}ai?	+	kà	---->	s^{h}ai?-ka	'trishaw'
	'side'		'car'		[< ENG: side car]	
	[< ENG]		[< ENG]			
c.	gai?	+	tañ	---->	gai?-tañ	'yardstick'
	'yard'		'stick'			
	[< Hindi]		[native BS]			

The second source of voiced obstruents in Burmese are aspirated obstruents. Note however that, as seen in (12), aspirated sonorants such as *hm, hmy, hn, hny, hŋ, hl, hy* are not voiceable, and remain unchanged in contexts that otherwise produce junctural voicing (cf. Okell 1969: 13).

(12) Aspirated Sonorants

a.	θà	+	hmwè	---->	θà-hmwè	'animal fur'
	'animal'		'hair'			
b.	koù	+	hni?	---->	koù-hni?	'nine years'
	'nine'		'car'			
c.	nwà	+	hlè	---->	nwà-hlè	'bullock cart'
	'cow'		'cart'			

In contrast to aspirated sonorants, as seen in (13), aspirated obstruents can undergo junctural voicing.

(13) Aspirated Obstruents

a.	sà 'eat'	+	phà 'frog'	---->	sà-**bà**	'seasoned opportunist'
b.	kà 'picture'	+	chaʔ 'flat thing'	---->	kà-**ja**ʔ	'cardboard, chart'
c.	s^hèi 'medicine'	+	khàn 'room'	---->	s^heì-**gà**n	'clinic'
d.	pè 'peanut'	+	shi 'oil'	---->	pè-**zi**	'peanut oil'
e.	sà 'eat'	+	khweʔ 'cup'	---->	sà-**gwe**ʔ	'dishes'

Burmese has another voicing alternation phenomenon which I refer to as "parasitic voicing". In order to understand this phenomenon, it is necessary to first understand Burmese rhyme reduction.

In Burmese compounding, as seen in (14), the rhyme of the first syllable can reduce, and this reduced vowel may be represented as [ə].

(14) Rhyme reduction

a.	phà 'frog'	+	pyañ 'fly'	---->	phə-**by**añ	'tree frog species'
b.	phà 'frog'	+	pyouʔ 'boil'	---->	phə-**by**ouʔ	'toad'
c.	θu 'person'	+	cì 'be big'	---->	θə-**j**ì	'chief'

Note that rhyme reduction in (14) is not because of voicing on the initial consonant of the following syllable. In other words, rhyme reduction occurs independently of junctural voicing, otherwise, there could not be rhyme reduction in (15) where the second syllable is already a voiced nasal.

(15)a.	pà 'cheek'	+	moú 'mound'	---->	pə-moú 'prominence of cheek'

b.	pà + 'cheek'	mòuñ + 'hate'	hmwè	---->	pə-mòuñ-hmwè	'sideburns'
			'hair'			
c.	θà +	nañ 'flesh'	+ yi 'smell'	---->	θə-nañ-yi 'liquid'	'pus'
d.	ti? + 'one'	naiñ 'win/can'		---->	tə-naiñ	'work within one's capacity'

In fact, as seen in (16), rhyme reduction (sometimes even onset reduction) occurs even in disyllabic compounds where both consituents have nasal initials.

(16) Rhyme reduction in disyllabic compounds

a.	nwà + 'cow'	má 'female'	---->	nə-má	'cow (female)'
b.	hna + 'nose'	màuñ 'lever'	---->	hnə-màuñ	'trunk (of an elephant)'

In addition, as seen in (17), rhyme reduction occurs when the second syllable is aspirated sonorants which remain unchanged in contexts that otherwise produce junctural voicing.

(17) Rhyme reduction with voiceless sonorants as second syllable

a.	thàñ + 'toddy'	hn̥a? 'clamp'	---->	thə-hn̥a?	'wooden clamp'
b.	s^{h}añ + 'hair'	hnei? 'press'	---->	s^{h}ə-hnei?	'hair clip'

Although the reduced syllable *Cə* loses its tonal contrast and rhyme quality, it maintains its syllabicity in the compound. As seen in (18), number one *ti?* is always reduced when it precedes a classifier.

(18) Reduced number one

a.	ti? + 'one'	pa? 'round'	---->	tə-ba?	'one round'
b.	ti? + 'one'	ma? 'quarter'	---->	tə-ma?	'one fourth'

c.	tiʔ +	hlwà	---->	tə-hlwà	'the whole'
	'one'	'range'			

But as shown in (19), this number does not reduce when preceding other numbers.

(19) Non-reduced number one[5]

s^he +	tiʔ	---->	s^hé-tiʔ	'eleven'
'ten'	'one'			

In cases where the first syllable reduces to *Cə*, junctural voicing may be extended leftward to the onset of the first syllable (cf. 20). I refer to such cases as parasitic voicing, a "non-local" process in the sense that consonant voice assimilation skips over a vowel.

(20) Parasitic Voicing

a.	s^hañ +	thoù	---->	**zə-doù**	'hairpin'
	'hair'	'pinch'			
b.	s^hañ +	kha	---->	**zə-ga**	'sieve'
	'rice'	'shake'			
c.	θwà +	teʔ	---->	**ðə-deʔ**	'canine tooth'
	'tooth'	'rise'			
d.	sà +	pwè	---->	**zə-bwe**	'table'
	'eat'	'feast'			
e.	tañ +	chu	---->	**də-ju**	'pole for picking fruit'
	'stick'	'pick (fruit)'			
f.	pyà +	tu	---->	**bə-du**	'hornet'
	'bee'	'hammer'			

As seen in (21 & 22) parasitic voicing is optional in some cases.

[5] There should be junctural voicing on the initial consonant of the second syllable in the compound *shé-tiʔ* 'eleven'. As Thurgood noted however, creaky tone can result from fusion of the genitive marker *ʔiʔ* to the possessive noun. In such cases, because the genitive marker *ʔiʔ* ends in a glottal stop, it subsequently prevents junctual voicing. Notably, *shé-tiʔ* 'eleven' is a result of *she* + *ʔiʔ* + *tiʔ* 'ten + of + one' (cf. Thurgood 1981:57).

(21)	a.	kàñ +	pà	---->	kə-**bà**	'bank of a river'
		'bank'	'vicinity'			
	b.			>	**gə-ba**	
(22)	a.	càñ +	poù	---->	cə-**boù**	'bed bug'
		'floor'	'insect'			
	b.			>	**jə-boù**	

In rare cases, as in (23 & 24), two degrees of reduction are attested in relation to parasitic voicing.

(23)	a.	thàñ	+	pin	---->	thə-**biñ**	'toddy palm'
		'toddy'		'tree'			
	b.				>	tə-**biñ**	
	c.				>	**də-biñ**	
(24)	a.	khà	+	pa?	---->	khə-**ba?**	'belt'
		'waist'		'go round'			
	b.				>	kə-**ba?**	
	c.				>	**gə-ba?**	

Note that in order for parasitic voicing to occur, voicing must be acquired on the initial consonant of the following syllable. As seen in (25), if the first syllable is reduced in Burmese compounding, aspirated obstruents in the second syllable might not undergo voicing.

(25)		σ	+	asp. obst.	>	$\sigma_{\text{+ rime reduction}}$ - asp. obst.	
	a.	pà	+	chaíñ	---->	pə-chaíñ	'dimple'
		'cheek'		'crack'			
	b.	pà	+	chou?	---->	pə-chou?	'muzzle'
		'cheek'		'control'			
	c.	θu	+	khù	---->	θə-khù	'thief'
		'person'		'steal'			
	d.	sa	+	chi	---->	sə-chi	'secretary'
		'letter'		'lift'			
	e.	sà	+	phou	---->	sə-phou	'fireplace, kitchen'
		'eat'		'bellow'			

f.	sà	+	khwe?	---->	sə-khwe?	'trough, manger'
	'eat'		'cup'			

In the Burmese lexicon, there are some native head-entry words which begin with voiced consonants in modern Burmese (most voiced consonants are in loanwords in modern Burmese). The writing system [Written Burmese (WB)] shows that those native voiced consonants are very likely the result of reanalyses, i.e. the second syllable of the compound is reanalyzed as an independent word. For example, the full Burmese word for 'head' is ʔù-gàuñ [< WB: ʔù-khàuŋ], which literally means 'head-hollow' (cf. 26). In this two syllable compound, the first syllable does not reduce to Cə, which mean that there would be voicing assimilation on the initial consonant of the second syllable (cf. 13). It appears that over time the Burmese reanalyzed the second voiced consonant (which was the result of junctural voicing) as an independent word (cf. Okell 1969:114).

(26) Reanalysis (cf. Junctural voicing rule in (13))

ʔù-khaùŋ > ʔù-gàuñ > **g**àuñ MED:61

As mentioned above (cf.12), aspirated sonorants are not voiceable, and therefore, remain unchanged in contexts that otherwise produce voicing. As seen in (27), there are a few exceptions to this rule.

(27) Exceptions

a. hnàñ +	pha? +	chin	---->	**nə-ba**?-chin	MED:247
'sesame'	'substance'	'sour'		'sesame oil-cake'	
b. kou +	sà +	hle	---->	kou-**zə-le**	MED:25
'body'	'instead'	'exchange'		'delegate'	
c. hŋe? +	pyò		---->	**ŋə-by**ò	MED:100
'bird'	'long object'			'banana'	

Conclusion

Intervocalic voicing assimilation occurs in Burmese compounding, reduplication, and with grammatical markers, which I refer to as junctural voicing. If the second syllable's onset is a voiceless obstruent, junctural voicing occurs (cf. 5). In cases where the second syllable begins with an aspirated obstruent, if the first syllable does not reduce to [ə], voicing will appear in the onset of the second syllable (cf. 13).

In cases where the first syllable reduces to *Cə*, junctural voicing may spread leftward to the onset of the first syllable. I refer to such cases as parasitic voicing, a "non-local" process in the sense that consonant voice assimilation skips over a vowel. As schematized in (28), parasitic voicing occurs optionally. It appears that this alternation is somewhat lexicalized, and is not predictable based upon some natural class.

(28) Parasitic Voicing

a. Voiceless Obstruents
vl. obst. + voiceable obst. > vl./**vd**$_{[+\text{ reduction}]}$ + **vd**. obst. (cf. 21 &22)

b. Voiceless Aspirates
vl. asp. + voiceable obst. > asp./vl./**vd**$_{[+\text{reduction}]}$ + **vd**. obst. (cf. 23 &24)

Parasitic voicing seems to happen because the minor syllable of the resulting sesquisyllabic structure (Matisoff 1973) is so reduced that voicing is allowed to spread leftward, on condition that voicing has occurred (or has been acquired) on the second syllable. Note that, de-aspiration occurs optionally in the onset of the first syllable in parasitic voicing (cf. 24b & 25b).

It appears that the language is acquiring some phrase-initial phonemic voiced consonants in its lexicon by way of reanalysis.

References

Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages.* Cambridge: Blackwell.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. Larry M. Hyman ed. *Consonant Types and Tones*. Los Angeles: University of SC.

______.1976. Introduction to the Written Burmese Rhyming Dictionary. *LTBA. Vol.3:1.*

Myanmar Language Commission. 1994. *Myanmar-English Dictionary* (MED). Yangon: Ministry of Education.

Okell, John. 1969. *A Reference Grammar of Colloquail Burmese. Two Vol.* London: Oxford Univ. Press.

Roop, Haigh D. 1972. *An Introduction to the Burmese Writing System*. New Haven: Yale University Press.

Soe, Myint. 1999. *A Grammar of Burmese*. Ph.D. Diss. University of Oregon, Eugene, OR. (MS).

Thurgood, Graham. 1981. *Notes on the Origins of Burmese Creaky Tone.* Tokyo: Institute for the study of languages and cultures of Asia and Africa.

Parasession

Phonetic Sources of Phonological Patterns: Synchronic and Diachronic Explanations

Phonetics and Phonology of Transparent Vowels in Hungarian

STEFAN BENUS[†], ADAMANTIOS GAFOS[‡], LOUIS GOLDSTEIN[¶]
[† ‡] *New York University*, [‡ ¶] *Haskins Laboratories*, [¶] *YaleUniversity*

1. Introduction

Vowel harmony is a requirement by which vowels in a certain domain agree in one or more phonetic features.[1] In Hungarian, the feature subject to harmony is the horizontal position of the tongue ([±back]). In many Hungarian roots, vowels in a word are either all front or all back, as in *öröm* 'joy', *város* 'city' (umlaut denotes front round vowels, and acute accent denotes length). In such roots, the backness of the suffix vowel is determined by the backness of the root vowels, e.g. *öröm-nek* 'joy'(dative), *város-nak* 'city'(dative).

In the so-called disharmonic roots, front vowels can combine with back vowels. In these roots, the quality of the root-final vowel determines the quality of the suffix vowel as in *parfüm-nek* 'perfume'(dative), *nüansz-nak* 'nuance'(dative). Vowels such as /ü/, /a/ are called opaque, because they block agreement between the initial and the suffix vowels. Certain vowels, however, do not block such agreement: *papír-nak* 'paper'(dative), *kávé-nak* 'coffee'(dative). Vowels like /*í*/ and /*é*/ are called transparent (henceforth TV).

Despite the significant body of work on vowel harmony (Clements 1977, S. Anderson 1980, Kiparsky 1981, Hulst & Smith 1986, Archangeli & Pulleyblank 1994, Ohala 1994, Ní Chiosáin & Padgett 1997, Ringen & Vago 1998, Baković & Wilson 2000, Krämer 2001), surprisingly little attention has been devoted to the phonetics of TVs. In line with the current research program on the role of phonetics in phonology (e.g. Steriade 1997), we believe that in order to understand the nature of transparency, both phonetic and phonological data should be studied. In this paper, we report preliminary results from such a study. At a broad level, we argue that phonetic details play a crucial role in determining the nature of transparency as well as the phonological pattern of suffix selection.

2. Transparent vowels

The Hungarian TVs consist of the front unround vowels {i [i], í [i:], é [e:], e [ɛ]}. In monosyllabic roots, TVs usually select front suffixes, as shown in (1). This is

[1] Work supported by NIH Grant HD-01994 to Haskins Laboratories.

expected since TVs are front. However, there is a limited set of about sixty mostly monosyllabic roots where {í, i, é} select back suffixes, as in (2) (Vago 1980).

(1)	cím-nek/*nak	'address'(dative)	szél-nek/*nak	'wind'(dative)
	hisz-nek/*nak	'believe'(3rd p. pl.)	szem-nek/*nak	'eye' (dative)
(2)	híd-nak/*nek	'bridge'(dative)	cél-nak/*nek	'aim'(dative)
	nyit-nak/*nek	'open'(3rd p. pl.)		

Not all TVs are equal, however. In polysyllabic roots, a systematic difference between {i, í, é} vs. /e/ emerges. If a back vowel (denoted with A) is followed by one of the three TVs {i, í, é}, the suffix vowel must be back, as shown in (3). If a back vowel is followed by /e/, suffix vowels vacillate between the front and back version, as shown in (4). Vacillation in suffix selection results also when the back vowel is followed by two TVs, as shown in (5). Finally, as in (6), when the second of the two TVs is /e/, the suffix vowel must be front.

(3)	A + {i, í, é}	papír-ban/*ben	'paper'(inessive)
	back suffix	buli-ban/*ben	'party'(inessive)
		kávé-ban/*ben	'coffee'(inessive)
(4)	A + /e/	hotel-ban/ben	'hotel'(inessive)
	vacillation	Ágnes-ban/ben	'Agnes'(inessive)
(5)	A+TV+TV	aszpirin-ban/ben	'aspirin'(inessive)
	vacillation	oxigén-ban/ben	'oxygen'(inessive)
(6)	A+TV+/e/	kabinet-*ban/ben	'administration'(inessive)
	front suffix	november-*ban/ben	'november'(inessive)

Clearly, transparency is not a binary quality of vowels. Rather, the vowels {i, í, é, e} display a continuum of phonological behavior with full transparency and full opacity at the two extremes (compare 3, 6). Both the number and the quality of the TVs are factors affecting suffix selection (see also Farkas and Beddor 1987, Ringen and Kontra 1989).

3. Articulatory experiment

In the experiment, we tested the well-accepted assumption in phonology that TVs do not participate in vowel harmony (at least on the surface). We wanted to find out if the TVs are produced differently depending on their harmonic context ([a-i-a] vs. [e-i-e]). If there is no difference, the phonological assumption above would be supported also phonetically. If there is a difference, we wanted to find out if this difference is phonologically significant or the result of coarticulation.

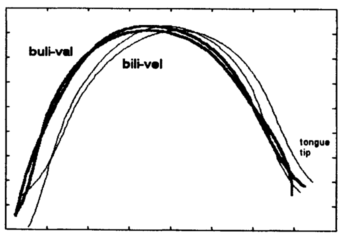
Four Hungarian subjects read a randomized list of stimuli words embedded in the frame sentence. Stimuli words contained TVs in front/back contexts with matched consonantal environment. A sample is given in (7).

(7)

Back	Front	Suffix
kábít-om [ka:bi:tom] 'daze'	répít-em [re:pitɛm] 'send'	1st p. sg. poss.
buli-val [bulival] 'party'	bili-vel [bilivɛl] 'pot'	Instrumental
bódé-tól [bo:de:to:l] 'hut'	bidé-től [bide:tø:l] 'bidet'	Ablative
hárem-ba [ha:rɛmbɔ]'harem'	érem-be [e:rɛmbɛ] 'medal'	Illative

We assumed that the participation of the TVs in palatal vowel harmony is directly related to the degree of horizontal tongue body retraction. Since all TVs are front, the target of the tongue body movement for a TV was assumed to be its extreme front position. To observe this movement, we first employed Ultrasound (Stone 1997). We traced the tongue surface at the target location of the TV.[2] Then we compared the obtained surface shapes from the same TV in a front and in a back context. We found that TVs are generally more retracted in the back harmony context. This is illustrated in (8).

(8) Retraction of /i/ in back vs. front harmony (*buli-val* vs. *bili-vel*)



In order to quantify the differences observed in Ultrasound, we employed Emma (Perkell et al. 1992). Emma offers highly precise kinematic data about the movement of a limited number of points (receivers) on the tongue. As with the Ultrasound data, we measured the most frontward position of the receiver on the tongue body during the TV. Then we calculated the difference in tongue body retraction between the front and back environments.

We report the results from one subject. Analysis of the other subjects is currently in progress. It was found that the horizontal position of the tongue body receiver differs as a function of the harmonic context: the TVs are more retracted in the back context than in the front context. The average difference between front and back contexts is 0.67mm. Pooling across vowels, the tongue body receiver locations show a significant difference (paired t-test, $p = 0.006$, $N = 232$ pairs). Per-vowel differences were /í/ 0.51; /i/ 0.06; /é/ 1.43; /e/ 0.72 mm.

Can the observed differences in tongue dorsum retraction be due to coarticulation from the adjacent vowels? To address this question, retraction effects must be compared between a vowel harmony context and a context where only coarticulation is in effect. In our stimuli set, we had two such comparisons.

[2] In our extractions, we employed a method developed by Khalil Iskarous at Haskins Laboratories.

First, note that vowel harmony in Hungarian is blocked across the boundary between two words, two parts of a compound, and between a verb and a prefix. A TV flanked by a harmony-blocking boundary from both sides may be retracted only due to coarticulation, see (9b). A TV in a harmony context, however, may be retracted due to coarticulation, harmony, or both, see (9a). Thus, a difference in retraction between TVs in (9a) and (9b) would demonstrate that the observed retraction in the back harmony context could not be due to coarticulation only.

(9)	a. **Harmony**	b. **Coarticulation**
	kabin-ból 'cubicle'(ellat.)	Jós[ka mit mon]d? 'What is J. saying?'
	kópé-ban 'prankster'(iness.)	Fic[kó B[be:] ban]da 'Group of Ficko B.'

Second, in some monosyllabic roots TVs select back suffixes, as in (10a), whereas in most such roots they select front suffixes, as in (10b). We compared TVs in these roots, as they appear in isolation (no overt inflectional markers). If any difference in tongue body retraction of the TVs between these two classes of roots is found, then it cannot be ascribed to coarticulation.

(10)	a. **Roots selecting back suffixes**	b. **Roots selecting front suffixes**
	síp 'whistle'(nom.)	cím 'address'(nom.)
	cél 'aim'(nom.)	szél 'wind'(nom.)

The results from these two conditions are not conclusive. We had 22 tokens of the data exemplified in (9) but did not find any significant difference. Therefore, the observed difference in the tongue body retraction between front and back harmony might be due to coarticulation only. However, prosodic differences between the forms in (9a) and (9b) make this a non-trivial comparison. For example, in the realization of the tokens by our subject, the syllable containing the TV was always unstressed in the harmony condition, but consistently received phrasal stress in the coarticulation condition.

Our stimuli included twelve pairs of the stems shown in (10). Data show a tendency for the TVs selecting back suffixes to be more retracted than the vowels selecting front suffixes. For example, tongue dorsum location differences between the two /í/s in the pair *síp-cím* were 1.77 and 1.19mm. This relationship obtains in nine out of twelve pairs. Unfortunately, the total number of pairs was small, which prevented statistical analysis. Despite this, the result from the monosyllabic roots suggests that the difference in tongue body retraction between front and back harmony context cannot be attributed to coarticulation.

To summarize the experiment, TVs are more retracted in the back harmony than in the front harmony context. This means that minute phonetic differences in tongue body position correlate with a full-fledged phonological alternation in suffixes. The data from bare roots suggest that this correlation cannot be explained by coarticulation only. Further experiments are necessary to obtain more robust results.

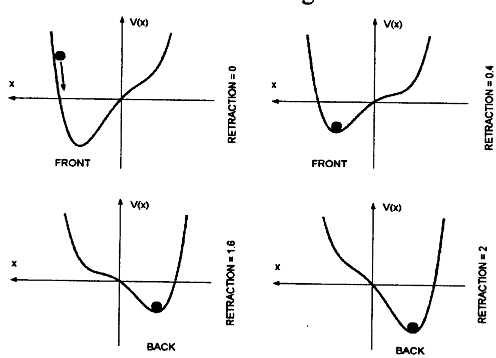
4. A dynamical model of suffix selection

How can small differences in details of articulation observed in section 3 be related to a categorical alternation in suffixes? Formally, the relation obtained between degree of retraction in the TV and suffix selection is nonlinear. Small changes in degree of retraction are associated with large (nonlinear) changes in suffix form. In this section, we propose a way to express this relation using the formal language of nonlinear dynamics (see Browman and Goldstein 1995 and Gafos, to appear, for background notions and motivation).

In the proposed dynamical model, the two discrete forms of an alternating suffix (e.g. dative *-nak* vs. *-nek*) are mapped to attractors of a dynamical system. We require that the choice of the attractor must be modulated by variation in R, representing the retraction degree of the root-final vowel. Mathematically, these ideas can be stated in the form of equation $dx/dt = N(x, R) + F(t)$. This equation expresses the temporal evolution of the suffix vowel (tongue body) constriction location variable, denoted by x, as a nonlinear function N of the current state x and the control parameter of retraction degree R. Given our requirements, a good candidate for N is the 'tilted' anharmonic oscillator, whose dynamics are described by $N(x, R) = (R-1) + x - x^3$. The factor F(t) represents the presence of noise and can be ignored for now.

The asymptotic behavior of x in this equation can be visualized by looking at (11). The abscissa denotes values of tongue body constriction location for vowels. Negative x values represent front vowels; that is, as x increases, the constriction location moves toward more retracted (back) locations in the vocal tract.

(11) Suffix form as a function of retraction degree



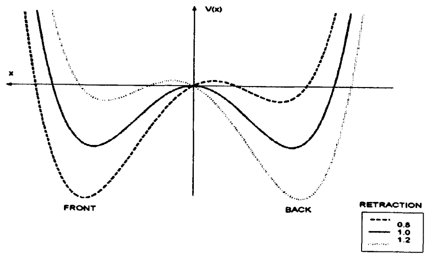
The value of constriction location for a suffix vowel can be interpreted by the position of a ball running downhill within the potential landscape $V(x) = (1-R)x - (1/2)x^2 + (1/4)x^4$, which can be obtained by integrating N(x, R). The minima in the valleys of V(x) represent the stable fixed points of constriction location. These are the attractors, the preferred regions within the continuum of constriction location where the ball ends up. Starting with a retraction degree of R = 0, the potential V(x) is shown at the left side of the top panel in (11). There is one attractor, FRONT, corresponding to the front variant of the suffix. A ball left in this potential will end up in that attractor, and the suffix surfaces as front. Small increases in R result in smooth changes in the potential V(x), as shown in the top, right where R = 0.4. There is still only one attractor or one possible stable form for the suffix. The graphs in the lower panel show how the potential changes after a significant increase in R, R = 1.6 left, R = 2 right (intermediate R values to be discussed). A qualitative change is evident in the shape of V(x). The FRONT attractor has been replaced by a BACK attractor that is located at the other end of the constriction location axis and corresponds to the back variant of the suffix.

To illustrate how this model captures the essential properties of the data, consider a root with a TV, e.g. *papír-nak* 'paper'(dative). Vowel gestures in consecutive syllables are overlapped (Öhman 1966). We assume that the constraint driving palatal harmony requires that the back vowel /a/ and the front vowel /í/ minimize their differences in (horizontal) tongue body constriction location. Consequently, /í/ must be retracted when following a back root vowel and this is what we observed experimentally. Let us set /í/'s retraction degree (when it follows a back vowel) to some relatively high value of retraction, say $R_í$ = 1.6 or higher. According to the dynamical model of the suffix alternation, this value of retraction induces the back variant of the suffix, as shown in the bottom panel of (11). In this model, then, a categorical alternation in suffix form is brought by a scalar increase in the continuous variable of retraction degree R in the preceding root vowel. This corresponds to the standard view of harmony as agreement between root vowels and the suffix. A crucial difference, however, is that in this model small, non-contrastive differences in retraction R can result in categorical suffix alternations (the property of nonlinearity).

Next consider the case of an opaque vowel, e.g. *parfüm-nek* 'perfume' (dative). The suffix vowel here surfaces as front. What makes /ü/ different from /í/ in this respect? We view transparency as grounded in the quantal nature of the relation between articulation and sound (Gafos 1999). On the one hand, Stevens (1989) and Wood (1979) have shown that the acoustic outputs for non-low front vowels are insensitive to a limited amount of variation in the horizontal position of the tongue body. Therefore, /í/ may be retracted to some degree $R_í$ without losing perceptual identity. On the other hand, Wood (1986) has shown that the front round vowels are less stable than their unround counterparts in that their acoustic output is very sensitive to even small amounts of articulatory variation in the horizontal position of the tongue. Consequently, if /ü/ were to be retracted to a degree comparable to $R_í$ it would lose its perceptual identity. Going back to our dynamical model in (11), then, if /ü/'s retraction degrees are limited to relatively small values ($R_ü < 0.4$), the attractor is FRONT and the suffix takes its front variant.

So far we have described the effects of varying retraction degree for relatively small and relatively large retraction values, as shown in (11). The result is a qualitative change from a front to a back attractor. In nonlinear dynamics, a change from one macroscopic state of the system to another implies an intermediate stage of fluctuation. This turns out to capture another essential aspect of the data, namely, vacillation. In (12), the potential V(x) is shown for three representative, intermediate values of R, 0.8, 1.0, 1.2. For each of these, there are now two minima representing the presence of two stable states, FRONT and BACK.

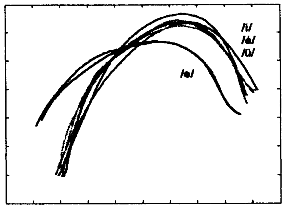
(12) Bistability for intermediate values of retraction



To see how bistability in an attractor landscape implies vacillation, we must turn to the effects of noise. In any model of a natural phenomenon, noise is introduced by the various low-level microscopic systems implementing the essential variables under modeling (here, the neuronal and myodynamic systems implementing tongue body constriction location). Mathematically, noise enters the model equation as a small, random fluctuation force F(t). This force pushes the position of the ball back and forth randomly. Consider, for example, a ball at position (0,0) in the bistable potential corresponding to R = 1 (solid line in 12). Due to the random kicks introduced by fluctuations, the ball will end up either at the left or the right side, and the suffix varies between a front and a back version.

There are two areas of vacillation in the Hungarian data. The first is vacillation induced by the low vowel /e/ as described in (4), e.g. *hotel-ban/ben* 'hotel'(inessive). The vowel /e/ then appears to be less transparent than {i, í, é}. Consistent with the Hungarian facts, in a cross-linguistic study of transparency, L. Anderson (1980) observes an implicational generalization related to vowel height: if /e/ is transparent, /i/ must be also but not vice versa. We have proposed that the transparency of {i, í, é} is grounded in their quantal properties, and specifically in the perceptual insensitivity of these vowels to articulatory perturbations in tongue body constriction location. Crucially, this is a property of the non-low front vowels (Stevens 1989, Wood 1979). Our ultrasound imaging showed that /e/ is different in a relevant sense from the other three TVs {i, í, é}. As shown in (13), /e/ is notably lower and more retracted compared to the vowels {i, í, é}, among which only minimal differences in height and backness can be observed (vowels were extracted from identical contexts). Because /e/ is less stable than {i, í, é}, it follows that /e/ can be retracted less than /í/, $R_e < R_í$. We may assume then that R_e falls within the range of intermediate values shown in (12). For such values, the dynamics of suffix selection has two attractors, and the suffix form will show variation due to random fluctuations.

(13) Tongue shapes of /e/ [ɛ], /í/ [i:], /é/ [e:], /ü/ [y] (two tokens per vowel)



We now turn to the other area of vacillation, concerning multiple TVs as described earlier in (5). The generalization is that the more TVs between the back root vowel and the suffix vowel, the more likely it is that the suffix is front. To see how this vacillation can emerge from our model, consider a root with two TVs, e.g. *aszpirin* 'aspirin'. The crucial idea is that the degree of retraction on TVs should diminish with the distance between the initial back vowel of the stem and the TV. The farther the TV from the initial back vowel triggering harmony, the less retracted it will be. Thus, the first TV in *aszpirin* is predicted to be significantly retracted by some degree R_1 due to the harmony requirement, dictating minimization of the articulatory distance in constriction location between it and the first back vowel. The second TV in *aszpirin* is also predicted to be retracted because it is required to agree in constriction location with its preceding vowel. Clearly, however, less retraction R_2 is predicted on the second TV as its preceding TV is more advanced than a prototypical back vowel. If R_2 falls within the intermediate range of values generating a bistable potential, the macroscopic result is variation in the form of the suffix, *aszpirin-ban/ben*. We plan to test the empirically predicted values of retraction in a future study.

To summarize, we have proposed that vowel harmony is driven by articulatory agreement between overlapping vowel gestures, and that this agreement is constrained perceptually. Vowels differ with respect to their potential for agreement: /í/ is most retractable, /e/ is somewhat retractable, and /ü/ is minimally retractable. Articulatory-perceptual quantal relations are crucial in determining the degree to which a vowel can be retracted without losing its perceptual identity. Finally, we relate the continuous scale of retraction degree to the discrete suffix alternation using nonlinear dynamics, a formal language where continuity and discreteness coexist and interact within a unified framework.

5. OT formalism

In the previous section, we described vowel harmony as a phenomenon where articulatory agreement among vowels is constrained by perceptual considerations. In this section, we use Optimality Theory (Prince and Smolensky 1993) to show how the conflict between the pressures for articulatory agreement and perceptual faithfulness is resolved with constraint ranking. The crucial proposal is that faithfulness between the input and output forms is evaluated separately for perceptual and articulatory domains. This is motivated by the quantal relationship between articulation and perception (Stevens 1989, Wood 1979).

We propose three basic constraints. Harmony is construed as an articulatory process mandated by the AGREE constraints in (14). AGREE(TBCL)$_{Rt}$ requires that consecutive vowels have identical values of tongue body constriction location. Thus, with respect to this constraint, 'a-[i]' is less harmonic than 'a-[ɯ]'. AGREE(TBCL)$_{Rt\text{-}Suff}$ requires that the backness of the suffix vowel depend on the retraction degree of the root-final vowel (section 4). This constraint is violated if, for example, the retraction degree is R = 2, but the suffix is front (see 11). In conflict with the AGREE constraints are the faithfulness constraints in (15), (16). The first mandates perceptual and the second articulatory constancy.

(14) AGREE(TBCL)$_{Rt}$ – Consecutive root vowels minimize their difference (distance) in terms of TBCL (Tongue Body Constriction Location).
AGREE(TBCL)$_{Rt\text{-}Suff}$ – The TBCL value of the suffix vowel is determined by the retraction degree R of the preceding root vowel.

(15) IDENT$_{Perc}$(FRONT) – Corresponding vowel gestures in the input and output are perceived as front.

(16) IDENT$_{Art}$(TBCL) – Corresponding vowel gestures in the input and output have identical specifications for TBCL.

Tableau in (17) shows how transparency is formalized in OT. The input is a bisyllabic root where a back vowel is followed by one of the TVs {i, í, é}, e.g. *papír*. The candidates show the perceptual output on the top in '[ ]' and the articulatory output on the bottom in '{ }'. The degree of retraction is given by the value of R and illustrated with arrows, where two arrows mean more retraction than a single one.

(17) Transparency; TVs maximize articulatory agreement with the initial vowel without compromising perceptual identity

(a-i)$_{Rt}$(nVk)$_{Suff.}$ (pa-pír)$_{Rt}$-(nAk)$_{Suff.}$	IDENT$_{Perc}$ (FRONT)	AGREE (TBCL)$_{Rt}$	AGREE (TBCL)$_{Rt\text{-}Suff.}$	IDENT$_{Art}$ (TBCL)
a. a - [i] – e {i} R = 0		**!		
b. a - [ɯ] – a {⇉i} R = 2	*!			**
c. a - [i] – e {→i} R = 1.6		*	*!	*
d. a - [i] – a ☞ {→i} R = 1.6		*		*

Candidate (17a) is faithful to the input both articulatorily and perceptually. Due to the significant articulatory distance between the non-retracted TV and the initial vowel, AGREE(TBCL)$_{\text{Rt}}$ is violated twice (compare with (17d) where the TV is retracted, decreasing the articulatory distance from the initial vowel). To avoid violating AGREE(TBCL)$_{\text{Rt}}$, the /i/ in (17b) is significantly retracted to {⇉i}, exemplified with R = 2. This, however, results in a vowel that is perceptually not front, violating IDENT$_{\text{Perc}}$(FRONT). The /i/ in candidates (17c,d) is retracted less than (17b), {→i}, R = 1.6, and each candidate incurs one violation of IDENT$_{\text{Art}}$(TBCL). Due to this retraction, the articulatory agreement with the initial root vowel is better than in (17a) but worse than in (17b). Hence, both (17c) and (17d) receive one violation of AGREE(TBCL)$_{\text{Rt}}$. Moreover, this retraction of /i/ does not change its perceptual identity and (17c,d) do not violate IDENT$_{\text{Perc}}$(FRONT). Based on (11), the dynamical system selects a back suffix if R = 1.6. Candidate (17c), however, selects a front suffix in violation of AGREE(TBCL)$_{\text{Rt-Suff}}$. Given that (17d) is the output, the ranking is IDENT$_{\text{Perc}}$(FRONT) >> AGREE >> IDENT$_{\text{Art}}$(TBCL). AGREE must dominate IDENT$_{\text{Art}}$(TBCL) since the opposite ranking would favor (17a) over (17d). Similarly, IDENT$_{\text{Perc}}$(FRONT) must dominate AGREE since the opposite ranking would favor (17b) over (17d). The two AGREE constraints are not ranked, since they are never in conflict; they refer to different domains. Intuitively, tableau (17) expresses the idea that TVs can maximize articulatory agreement with the initial back vowel while preserving their perceptual identity.

Tableau (18) shows the OT formalism of opacity. In the input, a back vowel is followed by a front rounded vowel. The intended /ü/ in candidates (18a,b) is retracted in order to avoid multiple AGREE violations. Since any retraction of front rounded vowels significantly affects perception (Wood 1986), both candidates violate top-ranked IDENT$_{\text{Perc}}$(FRONT). Note that the degree of retraction (R = 1.6) allowed for /i/ is not allowed for /ü/. Candidates (18c,d) show no or minimal retraction (R = 0.4), hence the double violation of AGREE(TBCL)$_{\text{Rt}}$. Candidate (18c), however, also violates AGREE(TBCL)$_{\text{Rt-Suff}}$. This is because, when retraction is minimal, our dynamical model in (11) dictates the front version of the suffix.

(18) Opacity; perceptual constancy prevents significant articulatory retraction

(a-ü)$_{\text{Rt}}$-(V{a,e})$_{\text{Suff.}}$ (parfüm)$_{\text{Rt}}$(nVk)$_{\text{Suff.}}$	IDENT$_{\text{Perc}}$ (FRONT)	AGREE (TBCL)$_{\text{Rt}}$	AGREE (TBCL)$_{\text{Rt-Suff.}}$	IDENT$_{\text{Art}}$ (TBCL)
a. a - [ʉ] – e {→y} R = 1.6	*!			
b. a - [u] – a {⇉y} R = 2	*!			
c. a - [y] – a {y} R = 0.4		**	*!	
d. a - [y] – e ☞ {y} R = 0.4		**		

Recall from (2) that TVs in some monosyllabic roots select back suffixes, e.g. *híd-nak* 'bridge'(dative). This pattern is formalized in a way similar to (17). The

TVs in these roots are lexically-specified for some degree of retraction sufficient to induce a back vowel according to the model in (11), R > 1.6. Output *híd-nak* satisfies AGREE(TBCL)$_{\text{Rt-Suff}}$ and since it reproduces the lexically specified articulatory retraction, while still perceived as front, it is selected as the winner.

Finally, the phonological pattern of vacillation follows from the determined ranking too. Both output forms, *hotel-nak* and *hotel-nek* 'hotel' (dative), have a retracted vowel /e/. Due to the quantal nature of /e/, more retraction would violate IDENT$_{\text{Perc}}$(FRONT) whereas no retraction would cause a fatal violation of AGREE(TBCL)$_{\text{Rt}}$. Given the model described in (12) and the intermediate degree of retraction $R_e \cong 1$, both *hotel-nak* and *hotel-nek* would fare equally on AGREE (TBCL)$_{\text{Rt-Suff}}$. Since they are tied on all constraints, they are both possible outputs.

6. Conclusion

We argued that the phonological pattern of suffix selection in Hungarian palatal vowel harmony must include reference to non-contrastive phonetic distinctions in the degree of tongue body backness of the transparent vowels. We sketched a theoretical model that allows us to relate these continuous phonetic distinctions to the discrete phonemic alternation using the mathematics of nonlinear dynamics.

References

Anderson, Lloyd. 1980. Using asymmetrical and gradient data in the study of vowel harmony. In R. Vago (ed.) *Issues in Vowel Harmony*, 1-48. Amsterdam: John Benjamins.

Anderson, Stephen. 1980. Problems and perspectives in the description of vowel harmony. In R. Vago (ed.) *Issues in Vowel Harmony*, 271-340. Amsterdam: John Benjamins.

Archangeli, Diana and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambrdige, MA: MIT Press.

Baković, Eric and Colin Wilson. 2000. Transparency, strict locality and targeted constraints. In R. Billerey & D. Lillehaugen (eds.) *WCCFL 19 Proceedings*, 43-56. Somerville, MA: Cascadilla Press.

Browman, Catherine and Louis Goldstein. 1995. Dynamics and articulatory phonology. In T. van Gelder & B. Port (eds.) *Mind as Motion*, 175-193. Cambridge, MA: MIT Press.

Clements, George. 1977. Neutral vowels in Hungarian vowel harmony: An autosegmental interpretation. In J. Kegl, D. Nash, & A. Zaenen (eds.) *Proceedings of NELS 7*, 49-64.

Farkas, Donka and Patrice Beddor. 1987. Private and equipollent backness in Hungarian. In A. Bosch, B. Need & E. Schiller (eds.) *Papers from the Twenty-Third Annual Regional Meeting of the Chicago Linguistic Society* (CLS 23), Part II, 90-105. Chicago, IL: Chicago Linguistic Society.

Gafos, Adamantios. 1999. *The Articulatory Basis of Locality in Phonology*. New York: Garland.

Gafos, Adamantios. To appear. Dynamics in grammar: Comment on Ladd and Ernestus & Baayen. In L. Goldstein, D. Whalen, C. Best & S. Anderson (eds.) *Varieties of Phonological Competence (Laboratory Phonology 8)*. Mouton de Gruyter.

Hulst, Harry van der and Norval Smith. 1986. On neutral vowels. In K. Bogers, H. v.d. Hulst, & M. Mous (eds.) *The Phonological Representation of Suprasegmentals*, 233-279. Dordrecht: Forris.

Kiparsky, Paul. 1981. Vowel harmony. Ms., MIT.

Krämer, Martin. 2001. Vowel harmony and correspondence theory. PhD dissertation, Heinrich-Heine-Universität, Düsseldorf.

Ní Chiosáin, Maire and Jay Padgett. 1997. Markedness, segment realization, and locality of spreading. Report LRC-97-01, Linguistic Research Center, University of California, Santa Cruz. [ROA-188].

Ohala, John. 1994. Towards a universal, phonetically-based, theory of vowel harmony. *Proceedings of International Conference on Spoken Language Processing*, 491-494.

Öhman, Sven. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *JASA* 39, 151-168.

Perkell, Joseph, Marc Cohen, Mario Svirsky, Melanie Matthies, Iñaki Garabieta & Michel Jackson. 1992. Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *JASA* 92, 3078-3096.

Prince, Alan and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado.

Ringen, Catherine and Miklós Kontra. 1989. Hungarian neutral vowels. *Lingua* 78, 181-191.

Ringen, Catherine and Robert Vago. (1998). Hungarian vowel harmony in Optimality Theory. *Phonology* 15, 393-416.

Steriade, Donca. 1997. Phonetics and phonology: The case of laryngeal neutralization. Ms., UCLA.

Stevens, Kenneth. 1989. On the quantal nature of speech. *Journal of Phonetics* 17, 3-45.

Stone, Maureen. 1997. Laboratory techniques for investigating speech articulation. In J. Hardcastle & J. Laver (eds.) *The Handbook of Phonetic Sciences*, 11-32. Oxford: Blackwell Publishers.

Vago, Robert. 1980. *The Sound Pattern of Hungarian*. Washington: Georgetown University Press.

Wood, Sidney. 1979. A radiographic analysis of constriction location for vowels. *Journal of Phonetics* 7, 25-43.

Wood, Sidney. 1986. The acoustic significance of tongue, lip, and larynx maneuvers in rounded palatal vowels. *JASA* 80, 391-401.

Phonetics and Phonology of Transparent Vowels in Hungarian

Stefan Benus
NYU, Linguistics Dept.
719 Broadway, 4th fl.
New York, NY 10003

sb513@nyu.edu

Adamantios Gafos
NYU, Linguistics Dept.
719 Broadway, 4th fl.
New York, NY 10003

adamantios.gafos@nyu.edu

Louis Goldstein
Haskins Laboratories
270 Crown St.
New Haven, CT 06511

goldstein@haskins.yale.edu

Antigemination: Natural or Unnatural History?*

JULIETTE BLEVINS
University of California, Berkeley

1. Introduction.

Within Evolutionary Phonology (Blevins, to appear), recurrent sound patterns are argued to be a direct consequence of recurrent types of phonetically based sound change. Synchronic markedness constraints of structuralist, generativist, and optimality approaches are abandoned, and replaced, for the most part, with historical phonetic explanations which are independently necessary. Under the general Evolutionary Phonology approach, any recurrent sound pattern which does not lend itself to phonetic explanation is problematic.

Already, this framework has proven useful in identifying new phonetic explanations for well documented recurrent sound patterns and for distinguishing sound patterns with a natural history in phonetic substance from those with an unnatural history involving rule inversion, rule telescoping, analogy, or language contact. To take just one example, consider the phonetic typology of metathesis presented in Blevins & Garrett (1998, to appear). Given the recurrent nature of certain metathesis sound patterns, but not others, historical phonetic explanations grounded in perceptual ambiguity and coarticulation are proposed. Nasal-obstruent metathesis is not accounted for by these phonetic explanations. Nevertheless, nasal-obstruent metathesis is attested in at least seven East Cushitic languages. In this case, as detailed by Garrett and Blevins (to appear), a seeming exception to the phonetic typology of metathesis has an unnatural history rooted in pre-existing patterns of morphophonological alternation.[1] Nasal-obstruent metathesis then is a recurrent sound pattern with an unnatural history. Within Evolutionary Phonology, it is only by understanding the precise history of language change that sound patterns with natural and unnatural histories can be distinguished. The general Neogrammarian dichotomy of regular sound change and analogy is strengthened by giving further substance to both the phonetic sources of regular sound change, and the morphological and phonological bases of analogical change.

In this paper, I examine another recurrent sound pattern which is potentially problematic for the Evolutionary approach. The sound pattern in question is one in which phonological syncope rules are sometimes blocked from applying if their output would create a sequence of adjacent identical consonants. This recurrent

* I am grateful to Bob Blust, Andrew Garrett, and Tony Woodbury for helpful discussion of the Chamorro, Old Irish, and Yup'ik syncopes respectively.

[1] Other cases of nasal-obstruent metathesis with unnatural histories are summarized in Blevins and Garrett (to appear).

sound pattern was first characterized and analysed by McCarthy (1986), who referred to it as 'antigemination'.

One case of antigemination analysed by McCarthy (1986:220-21) is that found in Afar, an East Cushitic language of Ethiopia. The Afar data in (1) illustrates the general antigemination pattern. Unstressed pre-tonic vowels are lost from open syllables when preceded by open syllables (1a-h) unless the consonants flanking the targeted vowel are identical, in which case there is no syncope (1i-k). Unlike other instances of apparent rule-blockage, the constraint on syncope in Afar cannot be attributed to a general constraint against geminates. Geminates in Afar occur freely both within and across morphemes: *cammi* 'uncle', *aa**bb**-uk* 'hearing', *da**ff**e-**s-s**-a* 'she seated' (from /daffey-is-s-a/ 'sit-CAUS-she-IMPF').

(1) Afar unstressed vowel syncope (Bliese 1981:212-16; Barillot 2002)

$$V \rightarrow \varnothing \; [_{\text{word}}CVC\_\acute{C}V$$

<u>Stressed</u>	<u>unstressed</u>	gloss
a. xamíla	xamlí	'swamp grass' (acc./nom.-gen.)
b. cagára	cagrí	'scabies' (acc./nom.-gen.)
c. kaxánu	kaxní	'love' (acc./nom.-gen.)
d. digíray	digré	'let him play'/ 'he played'
e. gutúca	gutcé	'push(pl.)'/ 'he pushed'
f. barísay	barsé	'let him teach'/ 'he taught'
g. digibté	digbé	'she/he married'
h. wagerté	wagré	'she/he reconciled'
	<u>no syncope</u>	
i. xararté	xararé	'she/he burned'
j. dananté	danané	'she/he hurt'
k. walaltá	walalá	'she/he hurt'

McCarthy (1986) argues that antigemination is a consequence of the Obligatory Contour Principle which prohibits adjacent identical elements in phonological representations. Formerly a constraint on lexical representations, the Obligatory Contour Principle (OCP) is extended by McCarthy to exert an active influence on the mapping between underlying and phonological surface forms. A hypothetical Afar form like ***danné* (from *danané*) is blocked because it contains a sequence of adjacent identical *n*'s.

Odden (1988) presents serious theoretical and empirical criticisms of the OCP-based account of antigemination. First, he highlights weaknesses related to phonological representations and notions of adjacency. Within McCarthy's model, antigemination is predicted to apply to tautomorphemic derived C_iC_i sequences, but not to heteromorphemic sequences, since morphemes are claimed to define independent 'tiers'. Where antigemination is expected but not found morpheme-internally, McCarthy analyses the segments in question as long-distance geminates. Where antigemination is not expected but attested across morphemes, tier-conflation is claimed to apply prior to syncope. As Odden notes, the freedom to represent C_iVC_i sequences as long-distance geminates when necessary to allow syncope, combined with the freedom to order tier-conflation before syncope to derive intra-morphemic antigemination greatly weakens the predictive power of the model. An important empirical observation is that antigemination is not found in certain languages. Odden concludes that OCP is not a principle of universal grammar.

Within Evolutionary Phonology, where recurrent sound patterns are, for the most part, a reflection of recurrent phonetically based sound change, syncope is easily explained, but antigemination is problematic. Syncopating sound patterns like the Afar rule in (1) are widespread cross-linguistically, and numerous sound changes of precisely this type have been proposed for distinct language families. Thurneysen (1980:67) describes a general syncope sound change for Old Irish, and similar across-the-board syncope of unstressed vowels has occurred in the history of many Austronesian languages. In Chamorro, Proto-Malayo-Polynesian schwa, an extra-short vowel, was lost in the environment VC_CV (Blust 2000:88). A similar syncope rule is attested in the history of the Central Alaskan Yupik (Fortescue et al. 1994, Jacobson 1984) where Proto-Eskimo schwa has been lost from the second of a word-initial string of open syllables. Another example of schwa syncope in open weak (odd-numbered) syllables has occurred in the history of Munsee/Delaware, an Algonquian language (Goddard 1982). These syncopating sound changes are illustrated in (2i-iv), where proto-forms are Old Irish, Proto-Malayo-Polynesian, Proto-Eskimo and Proto-Algonquian respectively.

(2) Syncope as sound change

i. Old Irish	*V > ø /[C_0V́C_CV	*dilese > dilse 'ownership'
ii. Pre-Chamorro	*ə > ø /[C_0V́C_CV	*qaləjaw > atdaw 'sun'
iii. Yupik	*ə > ø /VC_CV	*aləqa(R) > ałqaq 'older sister' (CAY)
iv. Pre-Munsee	*ə > ø /[(C)_]$_{\text{Weak Syllable}}$	*wetehkwani > wtohwan 'branch'

The phonetic explanation for syncope rules like those in (2) is straightforward: short unstressed vowels range in pronunciation from hyperarticulated vowels with recognizable quality to hypoarticulated segments which lack any noticeable formant structure. These hypoarticulated tokens are easily reinterpreted by subsequent generations as consonant release, or zero. Syncope rules, then, have a clear natural history, which we will continue to understand better as our knowledge of speech production and perception deepens.

The problem, however, is that when we look at the potential phonetic precursors of phonological syncope rules, like those illustrated in (2), we do not see evidence of antigemination. On the contrary, McCarthy (1986) describes *phonetic* syncope effects in Odawa, Modern Hebrew, English, Japanese and Hooper Bay Chevak where unstressed vowels *may* be omitted between adjacent identical consonants. (In these and other examples, C° indicates a consonant with audible release.) He argues that it is precisely the phonetic character of these variations, illustrated in (3), which makes the OCP irrelevant. Since the OCP is a constraint on phonological representations, it has no effect on phonetic implementation rules.

(3) Problem I: phonetic syncope does not show antigemination (McCarthy 1986)

<u>Language/family</u>	<u>variant surface forms</u>
Odawa /Algonquian	tətanisi, t°tanisi,ttanisi'he stays for a while'
Modern Hebrew /Semitic	nadədu, nad°du, naddu 'he/they wandered'
English /Indo-European	sınənım, sın°nım, sınnım 'synonym'
Hooper Bay Chevak /Eskimo	ənəni, ən°ni 'house-loc.sg'

In the historical domain, there is also evidence that regular syncopating sound changes have occurred without respecting antigemination. For the Central Alaskan Yup'ik (CAY) and Munsee developments summarized in (2iii-iv), there are apparent cases where syncope has applied between identical consonants. Relevant data for Yup'ik is presented in (4i) and that for Munsee in (4ii). In (4i.a-c) the expected reflex of intervocalic Proto-Eskimo *ɣ is CAY ɣ as shown by the forms in (4i.d-f). Attested [x] or [xx] then reflects earlier geminate *ɣɣ derived via syncope. In (4ii), proto-forms with weak C_iə.C_i sequences show Munsee reflexes with vowel loss, followed by degemination. Unami maintains some adjacent identical consonants in these contexts (Goddard 1979).

(4) Problem II: syncope as regular sound change does not show evidence of antigemination

i.	Proto-Eskimo	Central Alaskan Yupik	gloss
a.	*iɣəɣaɣ-	ixaɣ-	'lean (against)'
b.	*pəɣəɣaR-	pəxxaR-	'stay up all night'
c.	*təɣəɣ-	təxə-	'be hard'
		təxxi-	'get hard'
d.	*aɣə-	aɣə-	'go (over or past)'
e.	*iɣə-	iɣə-	'swallow'
f.	*təɣu-	təɣu-	'take'
ii.	Proto-Algonquian	Munsee	
a.	*ne-ne:me	ne:m (< nne:m)	'I see (it)'
b.	*kə-kawi:	kawi (Unami kkawi)	'he sleeps'
c.	*pe-pak-	pake:w (<ppake:w)	'it is flat'

Just as phonetic perceptual and coarticulatory accounts of metathesis do not predict nasal-stop metathesis, so phonetic perceptual and coarticulatory accounts of syncope generally fail to predict antigemination (though see section 4). With no clear phonetic basis for antigemination effects, the Evolutionary approach leads us to consider alternative unnatural histories for this recurrent sound pattern. This paper is a first attempt at such a consideration, as well as a search for any cases of antigemination whose history is arguably phonetically natural.

In this paper, I observe strong correlations between antigemination and other sound patterns. As I illustrate in section 2, antigemination is attested in languages with pre-existing geminate-singleton contrasts and in languages with with degemination, but not elsewhere. After demonstrating these correlations, I attempt to explain precisely where and how antigemination arises. In the most common cases discussed in section 3, syncope between identical consonants appears to be blocked just in case its output would give rise to neutralization of a *paradigmatic* opposition. This subcase comes under the general heading of 'non-homophony', 'anti-identity', or 'paradigmatic contrast' effects. Here, antigemination is a composite of natural phonetically based syncopating sound change, and independent morphological effects. In section 4, antigemination in Central Alaskan Yup'ik is claimed to have a singularly natural history. One question which arises is why the effects seen Yup'ik are not found in other languages. Section 5 highlights distinct empirical predictions made by this account in contrast to others and briefly explore implications of this study for modern phonological theory. A strong prediction of this model is that phonetic antigemination effects will not be phonologized unless rule inversion, rule telescoping, language contact, or paradigmatic effects are involved.

2. Languages with and without antigemination.
A summary of languages with claimed antigemination effects from McCarthy (1986) is given in (5).

(5) Languages with antigemination effects (McCarthy 1986)

	Language	Family
I. Pre-existing C/C: contrast	Afar	Cushitic/Afro-Asiatic
	Iraqi Arabic	Semitic/Afro-Asiatic
	Damascene Arabic	Semitic/Afro-Asiatic
	Tunisian Arabic	Semitic/Afro-Asiatic
	Tiberian Hebrew	Semitic/Afro-Asiatic
	Central Alaskan Yup'ik	Eskimo/Eskimo-Aleut
II. Degemination	Tonkawa	Isolate (Texas, USA)
	Modern Hebrew	Semitic/Afro-Asiatic

The languages in (5) are organized into two classes. In class I languages, a geminate/non-geminate contrast exists, and, it turns out, is also reconstructable for relevant subgroups. In class II languages, there is no underlying length contrast, but there is evidence for degemination. Some languages in class I also show evidence of degemination in certain contexts (e.g. Iraqi Arabic, Tunisian Arabic, Damascene Arabic).

Note that while pre-existing length contrasts and degemination are sound patterns strongly associated with antigemination, they cannot be used to predict whether or not antigemination effects will be found. A sample of languages without antigemination effects is given in (6). Syncope in Hindi shows no antigemination effects (Odden 1988:465), but Hindi does have an underlying geminate/non-geminate contrast (6.I). In Piro (Matteson 1965, Lin 1997), on the other hand, where underlying geminate/non-geminate contrasts are absent, there is active degemination of obstruents, yet syncope also shows no antigemination effects (6.II).

(6) Languages without antigemination effects

	Language/Family
I. Pre-existing C/C: contrast	Hindi/Indo-European
	Akkadian/Semitic
	Hooper Bay Chevak/Eskimo
II. Degemination	Piro/Arawakan
III. Neither	Klamath/Lutuamian (see below)
	Outlier Polynesian, Mussau,
	Trukese, Iban/Austronesian (Blust 1990)

On the other hand, if a language *does not* have a robust length contrast and *lacks degemination*, then it *will not* exhibit antigemination. A language of this type is Klamath, a Lutuamian language of south-central Oregon (Barker 1964, 1965). Unlike the class I languages in (5), Klamath does not have a robust underlying length contrast (6.III).

In Klamath, only a small number of stems are analysed as containing geminates which contrast with singletons in the same environment, though on the basis of internal reconstruction, it is clear that these are of recent origin. And unlike the class II languages in (5), there is no evidence of general degemination in Klamath. With the exception of /...s-s.../ clusters, sequences of identical consonants which arise across morpheme boundaries surface without degeminating. Some examples are shown in (7).

In Klamath, as illustrated in (8), short vowels in open stem-initial syllables are lost when a prefix is added. This rule of syncope does not show antigemination effects. Compare *qoqa* 'puts on, wears a dress' in (8a), with the causative *hosqqa* 'puts a dress on someone' from /hVs-qoqa/ via stem vowel copy, and prefix-induced syncope. In this example, as in (8b), a surface geminate is derived via syncope. In (8c) the expected form is *snoc'c'e:y'a*, however, laryngeal neutralization takes *c'c'* to *cc'* (Blevins, 1993).

(7) Another language without antigemination: Klamath (Lutuamian, Oregon)

i. Pre-existing C/C: contrast? Highly limited.
ii. Degemination? No. (only for /ss/)

a. *telli:na* 'looks off the edge'	/tel-li:na/ [D:100]
b. *tilnne:ka* 'rolls into a hole'	/tiln-ne:ka/ [D:115]
c. *tinne:ka* 'sets (of sun)'	/tin-ne:ka/ [D:116]
d. *lt^{h}ewwa* 'grazes in a flat place'	/lt^{h}ew-wa/ [D:225]

(8) Klamath syncope

	Base	Syncope: V →ø / Prefix + [(C)C_CV...	
a.	qoqa	hosqqa	/hVs-qoqa/'puts on a dress/CAUS' [D:157]
b.	lalamna	sallamna	/sV-lalamna/'puts a round obj. on back/REFL' [D:205]
c.	c'oc'e:y'a	snocc'e:y'a	/snV-c'oc'e:y'a/'melts/CAUS' [D:95]

Though there is no general degemination process in Klamath, as illustrated by the forms in (7), the language does simplify derived /...s-s.../ clusters to [s] (Barker 1964:95), as illustrated in (9a-c).

(9) Degemination of /...s-s.../ clusters in Klamath /hVs-/ causatives

	Base	Degemination	gloss
a.	sle'a	hesle'a < /hVs-sle'a/	'sees/ CAUS'[D:373]
b.	s'aywakta	has'i:watka < /hVs-s'aywakta/	'knows/teaches' [D:342]
c.	sacaq^{h}wa	hascaq^{h}wa < /hVs-sacaq^{h}wa/	'wash (hands)/CAUS' [D:347]
	Compare:		
d.	q^{h}oyqa	hosq^{h}i:qa < /hVs-q^{h}oyqa/	'recognize'[D:323]
e.	t^{h}sin	hist^{h}san < /hVs-t^{h}sin/	'grow/CAUS' [D:414]
f.	c'ayalc^{h}n'a	hasc'yalcn'a < /hVs-c'ayalc^{h}n'a/	'backs up/ CAUS' [D:83]

Compare *sle'a* 'sees' with *hesl'a* 'causes someone to see' in (9a), from /hVs-sle'a/. The causative prefix /hVs-/ is *s*-final, as illustrated in (9d-f). In this case, the expected surface form is ***hessl'a*, but /ss/ degeminates to [s]. Given the suggested correlation between degemination and antigemination in the class II languages in (5), it is instructive to examine word forms in Klamath where syncope could give rise to /s-s/ sequences. Words of this type, with stem-initial /sVsV…/, are shown in (10). Hypothetical surface forms under distributive *(C)CV-* or reflexive/reciprocal *sV-* prefixation are shown in bold.

(10) Klamath sVs-initial stems

	Base	RED + Hypothetical syncope + degemination		
a.	sasalk'ya	/sa-sasalk'ya/ > /sa-ssalk'ya/ > **sasalk'ya**	'quarrels'	[D:349]
b.	sesadwi	/se-sesadwi/ >/sessadwi/ > **sesadwi**	'sells'	[D:359]
c.	sosannqa	/so-sosannqa/ > /so-ssannqa/ >**sosannqa**	'wrestles'	[D:383]

As illustrated in (10), if syncope and degemination apply, causative and reflexive-reciprocal surface forms are *identical* to the base forms from which they are morphologically derived. Antigemination is clearly not the answer to why there is no evidence for syncope in (10). As shown in (8), Klamath does not show evidence of antigemination in parallel forms where degemination is not applicable. Despite morphological attempts to produce prefixed forms in (10), the phonology conspires to produce forms which are identical to their non-derived counterparts. It is precisely this sort of striking collapse of paradigmatic contrasts which, I will argue, leads to apparent antigemination effects in the majority of the world's languages.

3. Unnatural history.

In the following sections, I review cases of antigemination reported in McCarthy (1986) and elsewhere. In each case, I provide evidence that the constraint against syncope between identical consonants has an unnatural history. In 3.1, languages show antigemination in limited morphological contexts which suggest paradigmatic anti-homophony constraints (cf. Yip 1998, Crosswhite 1999). In 3.2 synchronic syncope constitutes diachronic rule inversion.

3.1 Paradigm effects.

3.1.1 Tunisian Arabic. Tunisian Arabic as describe by Wise (1983:168-170) has a rule of vowel syncope which deletes unstressed high vowels from open syllables.[2] Tunisian Arabic also has conditioned degemination: geminate consonants which are not pre-vocalic are simplified to singletons. The syncope rule is given in (11) along with several examples of its application in (11ii). Notice that in (11ii), syncope can be seen as feeding degemination.

[2] Syncope is obligatory for high vowels, and optional for non-high vowels. In Arabic transcriptions, capitals write consonants with secondary pharyngealization.

(11) Tunisian Arabic unstressed vowel syncope: $\underset{[+\text{high}]}{V} \rightarrow \text{ø} \,/[\ldots\sigma.\ C\_CV(C)]_{Wd}$

i. conditions for syncope not met		ii. syncope applies (+ degemination)		
zarbijja	'rug'	zarbíiti	'my rug'	< /zarbíjjiiti/
quffa	'basket'	qúfti	'my basket'	< /qúffiti/
/fumm/	'mouth'	fumha	'her mouth'	< /fúmmiha/
zarbijjítna	'our rug'			
quffítna	'our basket'			
fummi	'my mouth'			

Though McCarthy (1986:241) suggests that regular antigemination effects are found in Tunisian Arabic, this is not the case. As noted by Wise (1983:169-70), in nouns, adjectives, and participles, syncope applies regularly, even when adjacent identical consonant sequences would be created. The examples in (12) illustrate the absence of antigemination in participles.

(12) Tunisian Arabic syncope without antigemination: participles

i. conditions for syncope not met	ii. syncope applies (+ degemination)
a. mitɣaʃʃiʃ 'angry (m.)'	mitɣaʃʃa 'angry (f.)' < /mitɣaʃʃiʃ-a/
b. mSammim 'determined (m.)'	mSamma 'determined (f.)'< /mSammim-a/

McCarthy's reference then, must be to forms within verbal paradigms. In fact, the only place where apparent antigemination effects are observed in Tunisian Arabic is in verbs of the form $CVC_iC_iVC_i$. When these verbs take a *-V* suffix, as shown in (13), the syncope rule operative in (11) and (12) does not apply, despite the fact that phonological environments are parallel to the participles in (12), where syncope does apply.

(13) Tunisian Arabic syncope with antigemination: inflected verbs

i. conditions for syncope not met		ii. no syncope + degemin. where expected	
a. mitɣaʃʃiʃ	'angry (m.)'	ɣaʃʃiʃu	'they angered' (**ɣaʃʃu)
b. /xaffif-/	'alleviate'	xaffifu	'they alleviated' (**xaffu)
c. /qarr-/	'decide'	qarraru	'they decided' (**qarru)

Why are syncope and antigemination blocked in inflected verbs in Tunisian Arabic, while applying freely elsewhere? As in the hypothetical Klamath derivations shown in (10), the combination of syncope and degemination results in paradigm collapse within the verb system. Many of the $CVC_iC_iVC_i$ verbs in question are causative forms of CVC_iC_i stems. Given this, the consequence of syncope + degemination is to essentially undo the templatic morphology associated with causative formation. Wise (1983:170) suggests this possibility, providing relevant examples:

> It is possible that the rules [of syncope and degemination: JB] are suspended in this environment to avoid clashes with verbs with a CVCC stem; most CVCCVC verbs are in fact derived morphologically from verbs of CCVC or CVCC type and are therefore quite closely related to them semantically. For example, we find both /qarru/ 'they admitted' and /ɣaʃʃu/ 'they cheated' from /qarr+u/ and /ɣaʃʃ+u/.

In other words, what appears to block phonological rule application in (13ii) is not a phonological constraint, but a morphological one. The forms in (12ii.a) and (13ii.a) provide a minimal pair: in participles, the rules apply, but in verbs, they do not. Within the participle paradigm, verb stems of the form CCVC and CVCC have CVVCVC participles, with vowel length in the first syllable (e.g. *ʃaad* 'he pulled', *ʃaadid* 'pulling, have pulled'). As a result, there is no possibility of homophony arising via syncope between these members of the paradigm, as there is within the class of finite verbs.

The OCP-based account of antigemination in (13ii) can only be rescued by assuming that tier-conflation applies before suffixation of verbal inflection, but after suffixation of nominal inflection, or by equivalent domain restrictions on the OCP itself. This strictly phonological account relies heavily on the manipulation of tier-conflation (or its non-derivational equivalent), and misses the generalization true for both Tunisian Arabic and Klamath: regular phonological alternations may be blocked or unobservable precisely where they result in obliteration of paradigmatic contrasts.

3.1.2 Iraqi Arabic. In Iraqi Arabic as described by Erwin (1963:56-58), a syncope rule deletes short vowels from stem-final open syllables when vowel-initial suffixes are added. Iraqi Arabic also has conditioned degemination: as in Tunisian Arabic, geminate consonants which are not pre-vocalic undergo degemination. The syncope rule is given in (14) along with several examples of its application in (14ii).

(14) Iraqi Arabic unstressed vowel syncope: V $\rightarrow \emptyset$ /[...σ. C_CV(C)]$_{Wd}$

i. conditions for syncope not met		ii. syncope applies (+ degemination)		
ʃaʕar	'hair'	ʃaʕrak	'your hair'	< / ʃaʕar-ak/
xaabar	'he telephoned'	xaabrat	'she telephoned'	< /xaabar-at/
ybaddil	'he changes'	ybadluun	'they change'	< / ybaddil-uun/
ykassir	'he breaks'	ykasruun	'they break'	< /ykassir-uun/

Notice that, as in Tunisian Arabic, examples like *ybadluun* in (14ii) show syncope feeding degemination: /ybaddil-uun/ → /ybaddluun/ → *ybadluun*.

As in Tunisian Arabic, syncope without antigemination is attested. In (15), certain inflected adjectives undergo syncope giving rise to adjacent identical segments. Erwin (1963:242) makes it clear that syncope in this context is optional; nevertheless, it occurs, and leads one to question McCarthy's general claim that antigemination is a consequence of a general constraint on phonological representations.

(15) Iraqi Arabic syncope without antigemination: inflected adjectives

i. conditions for syncope not met	ii. syncope applies (optionally)
mitraaSiS 'crowded together m.'	mitraaSiSa, mitraaSSa 'crowded together f.'
	mitraSiSiin, mitraaSSiin 'crowded together pl'

However, syncope is consistently blocked in inflected verbs, whether they are associated with CVCCVC (class II) or CVVCVC (class III) templates, as shown in (16).

(16) Iraqi Arabic syncope with antigemination: inflected verbs

i. conditions for syncope not met		ii. no syncope + degemin. where expected
a. dallal	'he pampered'	dallilaw 'they pampered' (**dallaw)
b. jaddad	'he renewed'	jaddidaw 'they renewed' (**jaddaw)
c. ħaajaj	'he argued'	ħaajijaw 'they argued' (**ħaajjaw)
d. traaSSaS	'he moved close'	traaSSaSaw 'they moved close together' (**traaSSaw)

As in Tunisian Arabic, the failure of syncope to apply in (16ii) appears to be determined by a seeming anti-homophony constraint. If verbs of the form $CVC_iC_iVC_i$ or $CVVC_iVC_i$ undergo syncope and, in the first case, automatic degemination, under suffixation, they will collapse with inflected CVC_iC_i stems from which they are typically derived (Erwin 1963:65-66). For example Class II *tammam* 'to complete' is derived from /tamm/ 'to be complete'. A hypothetical form like ***tammaw* 'they completed' (< /tammam-aw/) is indistinguishable from *tammaw* 'they are complete. As in Tunisian Arabic then, the failure of syncope to apply between identical stem consonants in Iraqi verbs may better be explained in terms of paradigm collapse than OCP effects.

3.1.3 Damascene Arabic. The facts described for Damascene (Syrian) Arabic by Cowell (1964) are even more strikingly inconsistent with an OCP account than those noted above for other Arabic dialects. In Damascene, syncope optionally applies between adjacent identical consonants if they are both short (17d-eii), but fails to apply between identical consonants if they would produce an identical $C_iC_iC_i$ cluster (17iii).

(17) Damascene Arabic syncope: $\{e,o\} \rightarrow \emptyset$ /[...σ. C_CV(C)]$_{Wd}$

i. conditions for syncope not met		ii. syncope
a. btəskon	'you dwell'	btəskni 'you (f. sg.) dwell'
b. bisaaʕed	'he helps'	bisaaʕdu 'they help'
c. GalTet	'mistake of'	GalTTi 'my mistake'
d. biħaaʒeʒ	'to argue with'	biħaaʒʒu,'they argue with' (or biħaaʒəʒu)
e. ħaaTeT	'having put'	ħaaTTe 'having put, f.'
		iii. no syncope
f. bisabbeb	'he causes'	bisabbəbu'they cause' (**bisab(b)bu)
g. mSammem	'determined'	mSamməme 'determined, f.'

Precisely in this latter context, degemination is the norm. Notice that the morphological make-up of the cluster is not a factor: In examples like (17d) the identical sequence which results from syncope is tautomorphemic, while in (17e) syncope is blocked in tautomorphemic identical sequences.

In Damascene, as in the examples above, the underlying cause of antigemination effects appears to be potential paradigm obliteration. Since there is no surface contrast between $C_iC_iC_i$ and C_iC_i in the language (Cowell 1964:23-4, 27), in cases where syncope would give rise to $C_iC_iC_i$, the contrast between $CVC_iC_iVC_i$ and CVC_iC_i templates, so basic to the inflectional system of verbs, is potentially lost.

3.1.4 Tiberian Hebrew. McCarthy (1986) claims that Tiberian Hebrew syncope is blocked when a morpheme-internal geminate would be derived (18iii), but can give rise to adjacent identical consonants across morpheme boundaries (18ii.c).

(18) Tiberian Hebrew syncope: ə → ø /…σ. C_.CV

i. conditions for syncope not met	ii. syncope	
a. /zaaχər/	zaaχrúu	'they recalled'
b. /haaləχ/	haalχúu	'they walked'
c. /hinnen/	hinnii	'behold me' (< hinnnii hinnən-ii)
	iii. no syncope	
d. /Saaləl/	Saaləlúu 'they darkened'	
e. /saaβəβ/	saaβəβúu 'they surrounded'	

Odden (1988:467, footnote 11) emphasizes that McCarthy's entire argument for Tiberian Hebrew rests on the phonetic interpretation of the *shewa* symbol, which can be realized as ə or as nothing, and that the assumed contrast between, e.g. *zaaχərúu* (18a) and *saaβəβúu* (18e), occurs in one of the most controversial environments for interpreting the correct phonetic value of *shewa*. Relevant factors are shown in (19). With not clear basis for McCarthy's interpretation of *shewa* as schwa vs. zero in (18) , we are left to wonder whether the contrast between forms like those in (18ii) and (18iii) ever existed.

(19) Tiberian Hebrew *shewa* (Malone 1986, 1993; Odden 1988)
  a. phonetic value of *shewa* is *ə* or nothing
  b. grapheme *metheg* is thought to support interpretation of zero, but it is nonmandatory and therefore an unreliable indicator
  c. all cited examples of antigemination in Tiberian Hebrew contain *metheg*, irrespective of surrounding consonants

3.1.5 Tonkawa. Our knowledge of Tonkawa, an isolate of central Texas, is based on the collected works of Hoijer (1933, 1946, 1949, 1972). In Tonkawa, as described by Hoijer vowel syncope effects stem vowels which are non-final, preceded by at least one open syllable, and in open syllables themselves. Hoijer (1933:1) refers to stems as 'themes' and highlights the fact that vowel elision of this type is limited to themes, and never effects the form of affixes regardless of their position within the word. The syncope rule, stated in (20), is highly lexicalized then, applying only to stem vowels which themselves are both non-final in the stem and non-initial within the word.

(20) Tonkawa syncope: $\underset{[+stem]}{V} \rightarrow ø$ /V. C_.C $\underset{[+stem]}{V}$

Alternations which this rule is meant to capture are shown in (21), where abstract trivocalic stems /picena/ and /notoxo/ fail to surface with all three vowels intact.

(21) Syncope alternations in Tonkawa

a. picen	'steer, castrated one'	f. notox	'hoe'
b. picnoʔ	'he cuts it'	g. notxoʔ	'he hoes it'
c. wepcenoʔ	'he cuts them'	h. wentoxoʔ	'he hoes them'
d. picnanoʔ	'he is cutting it'	i. notxonoʔ	'he is hoeing it'
e. wepcenanoʔ	'he is cutting them'	j. wentoxonoʔ	'he is hoeing them'

Kisseberth (1970) notes that syncope is inhibited if a sequence of adjacent identical segments would be created. In (22), stems with regular syncope are compared with those in which syncope appears to be blocked. Though Kisseberth (1970) attributes this effect to the absence of underlying geminates in Tonkawa, McCarthy (1986) argues that antigemination in Tonkawa is another case where the OCP blocks syncope.

(22) Antigemination effects in Tonkawa

i. Underlying stem	ii. syncope		
a. /notoxo-/	notxoʔ	(< /notoxo-oʔ/)	'he hoes it'
b. /picena-/	picnoʔ	(< /picena-oʔ/)	'he cuts it'
c. /yakapa-/	yakpoʔ	(< /yakapa-oʔ/)	'he hits him'
d. /topo-/	ketpoʔ	(< /ke-topo-oʔ/)	'he cuts me'
	iii. no syncope		
e. /hewawa-/	he**wa**woʔ	(< /hewawa-oʔ/	'he is dead'
f. /ham'am'a-/	ham'**am**'oʔ	(< /ham'am'a-oʔ/	'he is burning'
g. /totopo-/	ke**to**topoʔ	(< /ke-to-topo-oʔ/)	'he cuts me rep.'

The reanalysis of antigemination proposed here, follows Hoijer's morphological analysis of Tonkawa closely. Of primary importance is the recognition that all sequences showing apparent antigemination effects are reduplicated C_iV_i sequences. A close examination of all examples of stem-internal C_iVC_i strings in Tonkawa, where V is a potential syncope target suggest that CV-reduplication is always involved. Reduplication in Tonkawa is used to productively mark repetitive aspect as well as plural subject or object, as shown by the pairs in (23) which are shown in their underlying (non-syncopated) forms.

(23) Productive CV-reduplication in Tonkawa (pre-syncope forms)

i.	<u>Non-repetitive</u>	<u>repetitive</u>	
	/topo-/	/totopo-/	'to cut (it) off'
	/nota-/	/nonota-/	'to touch'
	/kayce-/	/kakayce-/	'to be chopped off'
	/nawele-/	/nawewele-/	'to spread out a fabric'
	/tamaʔaxe-/	/tatamaʔaxe-/	'to be smashed, shattered'
	/notoxoko-/	/nototoxoko-/	'to expectorate'
	/yapece-/	/yayapece-/	'to sew, make clothes'
ii.	<u>non-plural subject</u>	<u>plural subject/object</u>	
	/nataya-/	/natataya-/	'to choose, select'

/panoxo-/	/panonoxo-/	'to bathe'
/napasxa-/	/napapasxa-/	'to play ball'
/noko-/	/nonoko-/	'to pick up'
/yatisxe-/	/yatitisxe-/	'to butt'

In some cases, reduplication holds of the first stem element like /to-to-po/ in (22g), while in other cases, it is the second stem element which is reduplicated as in /ha-m'a-m'a/ (22f). In some instances, reduplicated forms are somewhat semantically opaque, but synchronic evidence for a reduplicated base is still present. This is the case for /hewawa-/ 'die, be dead, be killed' (22e), with mediopassive prefix /he-/, to which we can compare /hehewa-/ 'to stop dying; recover (from an illness)' (Hoijer 1933:42-43), and /ham'am'a-/ 'to burn' (22f) with /ha-/ prefix (Hoijer 1933:39-42), and /m'a-/ theme also found in /m'a-ye-/ 'set fire to' (Hoijer 1946:301).

As noted by Hoijer (1933:7), reduplicated C_iV_i sequences have either the form $C_iV_iC_i$ or $C_iV_iC_iV_i$. In other words, no reduplicated string ever undergoes syncope of the first vowel of the derived C_iV_i-C_iV_i string, while some reduplicated strings will lose their second vowel, if it is in the context for syncope (20) to apply. In addition to the syncope rule in (20) then, reduplicated stems must be lexically marked as to whether their second vowel can or cannot undergo the syncope rule. Examples of these two reduplicated stem types are illustrated in (24). Notice that the failure of vowels to syncopate in (24ii) is entirely independent of phonotactics: when the stem is not reduplicated, syncope occurs between the same two consonants as shown in the plural object forms.

(24) Lexical syncope in reduplicated C_1V_1-C_2V_2 strings

i. Syncope of V_2		ii. No syncope of V_2	
yataso?s	'I stab him'	komo?s	'I suck it'
yaytaso?s	'I stab him rep.'	kokomo?s	'I suck it rep.'
/ya-yatasa-o?/		cf. wo**km**o?s	'I suck them'
coxno?s	'I sleep'	topo?s	'I cut it'
cocxa:yewo?	'several sleep together'	to**topo**?s	'I cut it rep.'
/co-coxa:-yewa-o?/		cf. we**tp**o?s	'I cut them'

What then accounts for the failure of syncope in forms like (22e-g)? I suggest that, as in the Arabic cases reviewed earlier, syncope is blocked just when its output would give rise to paradigm collapse. In this case, as in others, a regular degemination process is active in Tonkawa. As described by Hoijer (1946:292): "combinations of identical consonants always unite to form a single consonant." Examples of degemination across morpheme boundaries are illustrated in (25).

(25) Regular degemination in Tonkawa

<u>Underlying form</u>	<u>Surface form</u>	<u>gloss/text</u>
a. /tanmasla**k-k**^{w}a:low/	tanmasla**k**^{w}a:low	'jackrabbit' [T2.1]

b. /yako**n-n**acaka-.../ yako**n**acka:tewa:nes 'I'll knock you unconscious' [T1:4]

c. /ha:csokona**y-y**ayka'ay/ ha:csokona**y**ayka'ay-'a:la 'wolves-quot.' [T4:7]

Now consider the effect of syncope + degemination in the reduplicated forms where syncope appears to be blocked. Hypothetical derivations are in (26).

(26) Tonkawa C_1V_1-C_2V_2 strings, syncope, and degemination

<u>Base</u>	<u>RED with hypothetical syncope + degemination</u>			
a. /hewa-/	/hewawa-/ >	hewwa- >	****hewa-**	'to die'
b. /ham'a-/	/ham'am'a-/ >	ham'm'a >	****ham'a-**	'to burn'
c. /CV-topo-/	/CV-totopo-/	CV-ttopo-	****CV-topo-**	'to cut'

As with the Klamath data in (10), if syncope and degemination apply, repetitive and plural subject/object forms derived via reduplication will have exactly the same stem shape as the bases from which they are derived. Unlike Klamath, there are few if any forms outside of those involving reduplication where both syncope and degemination are applicable. However, an independent rule of stem-final vowel loss can feed degemination. As shown in (27), in compounds like those in (25), stem-final vowel loss gives rise to geminates which are subject to degemination.

(27) Final vowel loss (+ degemination): no antigemination in

a. /yako**na-n**acaka-.../ yako**n**acka:tewa:nes 'I'll knock you unconscious' [T1:4]

b. /yasʔe**ne-n**acaka-.../ yasʔenwencaka 'it was cutting them to death'[T19:11]

c. /yakexe-xakana-.../ /yakexakana-/ 'to push it down hard' [D:485.1]

As in the three Arabic dialects examined above, it is not a distinction between tauto- and heteromorphemic identical sequences which determines the distribution of antigemination effects. In productive reduplications like *totopo-*, where syncope is blocked, the CV reduplicative affix is arguably a distinct morpheme in Tonkawa. McCarthy's (1986) proposal that syncope is always blocked between tautomorphemic segments, but only sometimes blocked between heteromorphemic segments, is unsuccessful in predicting precisely where antigemination effects will occur. Antigemination in Tonkawa and the other languages examined above is not a general feature of sound patterns. It is in evidence precisely where vowel loss combined with other regular sound patterns would result in the phonological identity of two morpho-syntactically distinct forms within a paradigm. In the case of Tonkawa, syncope combined with degemination would result in merger of reduplicated and non-reduplicated stems which distinguish the inflectional paradigms in (23).

3.1.5. Modern Hebrew. McCarthy (1986) includes Modern Hebrew as a language with antigemination effects. The facts he cites in support of this are shown in (28).

(28) Modern Hebrew e/ø alternations in suffixed stems

i. conditions for syncope not met		ii. syncope applies	
a. kaʃar	'he tied'	kaʃru	'they tied'
b. kuʃar	'he was tied'	kuʃra	'she was tied'
c. hitkaʃer	'he contacted'	hitkaʃru	'they contacted'
		iii. no syncope	
d. nadad	'he wandered'	nadedu	'they wandered'
e. kucec	'he was cut'	kuceca	'she was cut'
f. titpalel	'I will pray'	titpaleli	'thou (f.) will pray'

In this case, Modern Hebrew differs quite dramatically from Arabic, and Tiberian Hebrew, where stems of the form shown in (28d-f) are always CVC_iC_i , - monosyllabic with a final geminate, when followed by a vowel. Under McCarthy's general analysis, antigemination was violated consistently in bilateral roots of this sort in the history of Arabic and Hebrew.

A clear alternative to McCarthy's interpretation of the Modern Hebrew pattern is that, geminates in Modern Hebrew have a peculiar unnatural history themselves. The absence of a length contrasts in the majority of Eastern European languages of the mid-nineteenth century, may have directly influenced the newly arising secular language. In this case, it could be that the historical geminate/non-geminate contrast was replaced with a C_i vs. C_ieC_i contrast directly. Since the syncope rule in question is not an innovation, the remapping of singleton/geminate contrasts can be seen to work on the output of the historical syncope rule. I have schematized the analysis in (29).

(29) Modern Hebrew reflections of historical C_i vs. C_iC_i contrasts

Mappings from the output of historical syncope:

Pre-Modern Hebrew		Modern Hebrew
$[\ldots C_i \ldots]_{stem}$	↔	$[\ldots C_i \ldots]_{stem}$
$[\ldots C_iC_i \ldots]_{stem}$	↔	$[\ldots C_ieC_i \ldots]_{stem}$

An alternative is to adopt an analysis parallel in nearly all respects to that made for Tonkawa above. If stems like *nadad* (28d) are viewed as partial reduplications of *nad-* , as suggested in Glinert (1989:460), then syncope plus degemination will result in neutralization of the paradigmatic contrast. Assuming regular degemination, what blocks syncope is the neutralization of distinct morphosyntactic paradigms.

3.2 Rule inversion and paradigm effects in East Cushitic.
Data from Afar, an East Cushitic language, was presented in (1). The Cushitic languages are distant cousins of the Semitic languages. In a very recent quantitative and comparative study of Somali, Afar, Rendille and Oromo, Barillot (2002) demonstrates that all of these languages show evidence of root OCP effects, and templatic non-concatenative morphology, well known from studies of their distant Semitic cousins (e.g. McCarthy 1981, 1982).

Within this context, Barillot (2002) also reviews the status of antigemination in these East Cushitic languages. One of his most significant findings in this area is that, in Somali, there is a correlation between syncope and identity of vowel quality between V_1 and V_2 in CV_1CV_2C stems. The general facts for Somali verbs are summarized as in (30). From a historical point of view, it is clear that the majority of CVCVC stems have the same vowel in both syllables because they are the result of vowel copy. The historical process is illustrated in (31), where Proto-East Cushitic is abbreviated PEC. Given historical evidence for vowel copy epenthesis, the generalizations in (30) follow from historical rule inversion: vowels which are historically present do not syncopate, while those which are historically inserted may take part in V/zero alternations.

(30) Salient characteristics of CVCVC verbs in Somali (Barillot 2002:279)

a. The majority of CVCVC stems have identical vowels in both syllables.

b. When CVCVC stems have two distinct vowels, *there is no syncope.*

(31) East Cushitic vowel copy: $*CV_iCC- > CV_iCV_iC- /\_\{C,\#\}$

a. PEC **bull-* 'flour'; Afar *bulul* 'become pulverized'; Oromo *bull-aw-* 'become pulverized'

b. PEC **ħizz-* 'yam, creeper'; Somali *ħidid*; Bayso *hidid*; Oromo *hidd-a.*

c. PEC **kilm-* 'tick'; Afar *kilim*, Oromo *ʃilm-a.*

However, in Somali, as in Afar, Rendille and Oromo, vowels are not lost between identical consonants in CVC_iVC_i stems even when vowels are identical. Barillot (2002) demonstrates that in some cases, this is because the medial consonant in Somali is (historically) geminate. For the remaining cases, he posits abstract underlying forms: stems like *barar-* 'enfler', he claims, are underlying quadrilateral $C_1VC_2C_1VC_2$ stems, with an empty onset to the second syllable (op cit. p.445). By assuming that these stems involve -CVC reduplication, Barillot at once accounts for the identity in vowels and consonants. The abstract medial empty onset is part of a consonant cluster and the structural description of the syncope rule is not met.

Barillot extends this abstract analysis to Afar, Rendille and Oromo where CVC_iVC_i stems also resist syncope, as illustrated in (32), from (Barillot 2002:465).

(32) East Cushitic antigemination effects in CVC_iVC_i verb stems

	Afar		Rendille		Oromo	
	se marier	*parler*	*avoir mal*	*se regrouper*	*traire*	*griffer*
1s/3ms	digba	**walala**	bolxa	**urura**	elma	**tarara**
2s/3fs	digibta	walalta	boloxta	ururta	elemta	tararta
1p	digibna	walalna	boloxna	ururra	elemna	tararna
2p	digibtan	walaltan	boloxtaan	ururtaan	elemtani	tarartani
3p	digban	**walalan**	bolxaan	**ururaan**	elmani	**tararani**
CAUS/	digbise	**walalise**	bolxica	**ururica**	elmama	**tararama**
PASS	*marier*	*faire parler*	*faire mal*	*regrouper*	*être trait*	*être griffé*

In (33) I summarize the descriptive facts for CVC_iVC_i stems which hold for all four East Cushitic languages.

(33) Salient characteristics of CVC_iVC_i verbs in East Cushitic
a. The second stem vowel does not alternate with zero
b. The final consonant is never guttural.
c. Vowels in the first and second syllable are identical.

Generalizations (33b,c) follow from the historical analysis presented in (31). Stems of the form CVC_iVC_i are derived from $*CV_iC_jC_j$- stems. It follows that consonants, like gutturals, which are not geminable, will not surface in these forms, and that the vowel in the second syllable will be a copy of the first, since vowel-copy epenthesis is involved. The generalization we need to account for now is that in (33a): why is it that all East Cushitic languages show antigemination in stems of this sort? Since the vowel/zero alternation originally involved vowel-insertion, not vowel deletion, there is no natural history of this process which involves a gradual evolution from phonetic to phonological syncope.

I suggest that in this case, rule inversion is coupled with clear paradigm uniformity effects. In all four East Cushitic languages under discussion, there is synchronic degemination in word-final and preconsonantal position, as in (34).

(34) East Cushitic degemination: $C_iC_i \rightarrow C_i$ / _{C, #}

As a consequence, under suffixation, CVC_iC_i- stems in the modern languages undergo regular degemination before consonant-initial suffixes, and word-finally. Some of these geminate final stems are lexical (Afar *obb-* 'hear'), but others appear to be the result of regular word-formation processes (e.g. Afar geminate-final imperatives). For the second class, stem-final gemination is morphological. In this case, paradigm leveling occur precisely where syncope would give rise to a stem-final geminate which is not morphological. As in Tonkawa, it is the morphological or morphotactic non-identity of two stem types (reduplicated and non-reduplicated in Tonkawa, final geminate vs. final C_iVC_i in East Cushitic) which is maintained under antigemination.

While this analysis involving paradigm leveling has a slightly different character from the direct anti-homophony effects seen in Arabic and Tonkawa, it is supported by facts from Arbore (Hayward 1984), where leveling has not occurred in CVC_iVC_i stem which reflect historical $*CVC_iC_i$-. In (35) relevant sub-paradigms are shown for three different stem-types: CVC_iC_i- stems, CVCVC- stems where the second vowel is non-alternating, and CVC(V)C- stems where the second vowel, in parentheses, alternates with zero in precisely the environments predicted by the historical reanalysis of (31) as vowel insertion.

(35) Some Arbore verb subparadigms (Hayward 1984:274-285).

	2s perfect	2s imperative	1s perfect	
i. CVCVC - stems				
ʔeruc- 'vomit'	ʔeructe	ʔeruc	ʔeruce	
tatab- 'want'	tatatte	tatap	tatabe	
kulil- 'warm oneself'	kulilte	kulil	**kulile**	} cf. bold forms in (35)
harar- 'hurry, be fast'	hararte	harar	**harare**	

ii. CVC_iC_i- stems			
horr- 'chase'	horte	hor	horre
fuss- 'miss'	fuste	fus	fusse
laww- 'milk'	lawte	law	lawwe
fayy- 'be saved'	fayte	fay	fayye
iii. CVC(V)C- stems			
zer(e)n- 'pierce'	zerente	zeren	zerne
war(a)b- 'fetch water'	waratte	warap	warbe
k'in(i)n- 'sting'	k'ininte	k'inin	**k'inne** } expected under rule inversion;
k'ad(a)d- 'shut'	k'adatte	k'adat	**k'adde** no antigemination

Note that CVC_iVC_i- stems which would violate antigemination under syncope fall into two lexical classes: those where the vowel is lost (violating antigemination), and those where the vowel is maintained, as in the related languages in (32). Since the Arbore system comes closest to representing the sound change with rule inversion, it supports an analysis in which the extension of non-alternating CVC_iVC_i- stems in other East Cushitic languages is due to analogical change.

4. Natural History.

In Blevins (to appear), I suggest that, despite the seemingly unnatural histories just examined, there is a potential natural history for antigemination effects. I propose that antigemination is a transient stage which occurs when syncopating sound changes innovate in languages with pre-existing geminate/singleton contrasts. The critical observation is that, at the stage of variation before sound change occurs, unstressed vowels in the syncope environment are reduced. If a reduced vowel can be re-interpreted as consonant release, or as a simple phonetic transition from one consonant to the next, then vowel loss is phonologized. However, in languages where underlying geminates occur, the audible release between identical consonants (or homorganic ones) in the syncope context will result in temporary resistance to their categorization by language learners as the same phonological entities as underlying (full or partial) geminates. These rearticulated identical elements will resist interpretation as geminates, because in true geminates, there is no release between the two 'halves' of the geminate consonant. At the same time, patterns of coarticulation will be expected to reduce or eliminate the open transition over time, so that eventually, despite the perceptual basis for resistance to syncope between geminates in such languages, if no other factors interfere, syncope should eventually occur. In all of the cases examined above, there is evidence that syncope can and will apply between identical consonants, unless paradigmatic constraints are operative. However, there is at least one known language where antigemination may perhaps be seen in its pure and natural state.

In Central Alaskan Yup'ik (CAY), with underlying geminate/non-geminate contrasts, schwa syncope applies across the board to unstressed vowels in the VC_CV context, unless the two consonants are identical (Miyaoka 1971, Reed et al. 1977, Woodbury 1982, Woodbury 1987, Woodbury personal communication, 2003). In this case, there is gemination of the post-schwa consonant, since in general schwa is not licit in unstressed open syllables. In Hooper Bay Chevak, the usual situation is for schwa to delete even when surrounded by identical consonants. Following Woodbury (1982), McCarthy (1986:245) is very explicit about the output of syncope in Hooper Bay Chevak: "The result of deleting schwa between identical consonants in no case merges with a true one-to-many geminate. Rather, the cluster of identical consonants is produced with a medial release that, in sonorant environments, is a full-fledged vowel...In no case does the derived cluster merge with a true geminate." Representative forms are in (36).

(36) Syncope and antigemination in Central Alaskan Yup'ik

	/kəmə-ni/	/kəmə-mi/	/ənə-ni/
General CAY	kəmni	**kəməmmi**	**ənənni**
Hooper Bay Chevak	kəmni	kəm˚mi	ən˚ni
	'his own flesh'	'of his own flesh'	'house-loc.sg'

McCarthy's account of the two differing dialects relies on representational differences which result from tier conflation. Morphemes are represented on separate tiers, which are conflated at some point in the derivation. If morphemes are on separate tiers, and syncope applies between identical consonants in different morphemes, the output of syncope will not violate the Obligatory Contour Principle, since adjacency is only defined tier-internally. The account is summarized in (37).

(37) An OCP-based account of Yup'ik dialect differences in vowel syncope

	antigemination	rule type	tier conflation
General CAY	yes	post-lexical	yes
Hooper Bay Chevak	no	late lexical	no

To my knowledge, no additional support for the distinction in rule type proposed by McCarthy exists, nor is there any evidence that syncope in Hooper Bay Chevak is anything but post-lexical. It is clearly not structure preserving, producing rearticulated adjacent identical consonants, and also follows other rhythmic rules which appear to be post-lexical (Woodbury 1982, 1987). McCarthy (1986:245) himself admits that "the distinction in the domain of syncope cannot as yet be independently motivated."

I suggest that the difference between General CAY and Chevak reflect the natural evolution of syncope rules. While pre-existing length contrasts and morphological effects may inhibit the progress of syncope between identical consonants, coarticulatory effects may eventually win out, as in the phonetic alternations illustrated earlier in (3). Central Alaskan Yup'ik stands out in the classification of languages presented in (5.I). In all other languages, templatic morphology defines whole paradigms or subparts of them. In all other languages, gemination can be the primary exponent of a morphological contrast, and this morphological contrast can, in effect, give rise to unnatural phonologization of the transient sound pattern defined by General Yup'ik.

Though the natural evolution of antigemination in Yup'ik appears to be unique cross-linguistically, this evolutionary stage (without Yupik-specific post-schwa gemination) is posited as a transient one for all phonetically natural syncopes in which a pre-existing geminate/non-geminate contrast exists. What is rare is for this transient state to be phonologized without the interference of factors external to phonetic naturalness. In all of the cases reviewed in section 3, a pre-existing morphological contrast appears to influence the grammaticization of antigemination. In Central Alaskan Yup'ik, the independent rule of post-schwa gemination (completely generalized in other dialects) may also have played a role: once unstressed $C_i ə C_i$ sequences are produced as $C_i ə C_i C_i$, syncope is further inhibited, since the schwa is no longer in an open syllable. Since Chevak lacks the post-schwa gemination rule, syncope is not further inhibited, and naturally extends itself to $C_i ə C_i$ sequences as well.

5. Explanation in Evolutionary Phonology.

The working hypothesis of Evolutionary Phonology is that common sound patterns typically result from common phonetically motivated sound change. Among these common sound patterns are the general syncope patterns examined above: in all but the East Cushitic case, synchronic syncope alternations mirror, to a great extent, phonetically natural syncopating sound changes. The specific question I have addressed in this study is whether the failure of syncope to apply between identical consonants in unrelated languages can also be viewed as phonologization of a phonetically natural effect. A survey of syncopating sound changes in which antigemination is not found suggests that non-phonetic processes are involved. All languages surveyed in section 3 show strong correlations between antigemination and homophony avoidance within paradigms. Nevertheless, in at least one language, Central Alaskan Yup'ik, antigemination shows no morphological conditioning and appears to have a near-natural history interrupted only by the existence of post-schwa gemination.

Does antigemination have a natural history? The answer within Evolutionary Phonology is complex: yes and no. Sequences of identical consonants which occur as variants of unstressed C_iVC_i sequences will more readily resist reinterpretation as geminates in languages with pre-existing consonantal length contrasts than in languages without them. If independent self-organizing principles of morphological analysis (i.e. morphological analogy) intervene during the course of language acquisition, antigemination may emerge as a consequence. However, where pre-existing length contrasts are absent, antigemination cannot have a natural history. If it is observed, as in Tonkawa, it is predicted to fall into the class of unnatural histories. Where it is not observed, as in the history of Munsee sketched in (4), the expected natural history of syncope continues uninterrupted.

How are the same facts, correlations and tendencies to be accounted for in purely synchronic terms? In sections 1 and 2 I outlined several problems for previous synchronic accounts of antigemination. McCarthy's (1986) approach is undermined by its dependence on multi-tiered representations (including long-distance geminates) and the operation of tier-conflation, both of which have been eliminated from standard Optimality-theoretic treatments (Gafos 1998, Keer 1999, Kager 1999). Odden's suggestion that the OCP is not a component of Universal Grammar leaves us in an even weaker position to understand the fundamental nature of the sound patterns in question.

The only well developed synchronic alternative to McCarthy (1986) I am aware of is the general account suggested by Rose (2000). Under her account, antigemination is still a phonological OCP effect, but the OCP operates only on C_iVC_i sequences, since any surface C_iC_i sequence in a given domain is claimed to be geminate and does not violate the OCP. Such geminates do, however, violate the NO-GEM constraint. The OCP and NO-GEM have distinct rankings in different languages, and McCarthy's tier-conflation is translated into domain-specific constraints. A serious empirical problem with Rose's account is her assumption that "any output sequence of two identical consonants within the same domain...constitutes a geminate, a single consonant with long duration. This is in line with phonetic evidence, which has found no distinction between surface true and fake geminates..." (Rose 2000:101). However, in McCarthy's original article, he notes that in languages like Modern Hebrew and Hooper Bay Chevak, precisely this contrast occurs: derived sequences of identical consonants can be rearticulated, but underlying geminates cannot.[3]

However, the fundamental problem I see with the range of synchronic approaches is not technical or empirical, but explanatory. While it is straightforward for synchronic approaches to reframe analyses of the antigemination cases described above in terms of the interaction of phonological constraints with morphological anti-homophony constraints (Yip 1998, Crosswhite 1999), or paradigm uniformity effects (Kenstowicz 1996, Benua 1997, Steriade 2000, McCarthy 2002, etc.), such accounts fail to make the same range of predictions as the historical approach advocated here. In (38) I highlight some of these predictions, and present them as a challenge for competing synchronic analyses.

The correlation in (38a) is a consequence of the unnatural histories involved in the majority of antigemination cases reported in the literature, as reviewed in this study. The phonological pattern ruled out in (38b) follows from our phonetic understanding of syncopating sound changes. Though certain variable $C_1{}^vC_2$, $C_1{}^\circ C_2$ sequences (where '^{v}' is a short unstressed vowel) may prove more resistant to (re)analysis as C_1C_2 than others, there is a strong phonetic tendency for adjacent identical consonants to merge into single segments over time. Only when non-phonetic analogical effects intervene, is there a recurrent pattern of grammaticized antigemination effects.

(38) Predictions of Evolutionary Phonology regarding syncope and antigemination

a. Antigemination is strongly correlated with languages which have either lexical geminate/non-geminate contrasts or degemination.
b. Pure antigemination as a regular feature of an exceptionless phonological syncope alternation with origins in unstressed/weak vowel loss is rare or non-existent. (General CAY is 'impure', due to post-schwa gemination).
c. In languages with only open syllables, production constraints may result in syncope between identical consonants only. (Blevins, 2003)

Finally, in (38c), I suggest a phonetic explanation for a recurrent pattern not discussed in this paper, but clearly related: in many languages with only open syllables, syncope occurs only between identical or homorganic consonants (Blust 1990; Odden 1988). Within synchronic accounts, the OCP and NO-GEM must both be low ranking, and an additional constraint must be invoked to rule out heterorganic sequences. The alternative is to note that in languages with only CV syllables, there is already an articulatory pattern that each consonantal feature complex must be released into a vowel. The only derived clusters which can be produced with this pattern intact are geminates or homorganic ones. There is no

[3] Similar facts are detailed for Imdlawn Tashlhiyt Berber by Dell and Elmedlaoui (1996).

reference to the OCP, NO-GEM, or any additional synchronic markedness constraints. The sound pattern is predicted by a simple interaction of syncope under pre-existing patterns of C-V coarticulation (Blevins 2003).

While the primary focus of this study is antigemination, I hope to have demonstrated more generally the extent to which Evolutionary Phonology provides a concrete model in which natural and unnatural sound patterns can be identified and, ultimately, understood.

References

Barillot, Xavier. 2002. *Morphophonologie gabaritique et information consonantique latente en Somali et dans les langues est-couchitiques.* Thèse de doctorat en linguistique théorique et formelle, Université Paris VII Denis Diderot, UFR do Linguistique.

Barker, M. A. R. 1963. *Klamath Dictionary.* Publications in Linguistics 31. Berkeley: Univ. of California Press.

Barker, M. A. R. 1964. *Klamath grammar.* UCPL 32, Berkeley and Los Angeles: University of California Press.

Benua, Laura. 1997. *Transderivational identity: phonological relations between words*. PhD dissertation. University of Massachusetts, Amherst.

Blevins, Juliette. 1993. Klamath laryngeal phonology. *International Journal of American Linguistics* 59: 237-279.

Blevins, Juliette. 2003. Constraints on syncope in Austronesian: perception, production, and sound change. Talk given at the Austronesian Circle, Honolulu, Hawaii. March, 2003.

Blevins, Juliette. To appear. *Evolutionary Phonology: The emergence of sound patterns*. Cambridge University Press.

Blevins, Juliette and Andrew Garrett. 1998. The origins of consonant-vowel metathesis. *Language* 74: 508-56.

Blevins, Juliette and Andrew Garrett. To appear. The evolution of metathesis. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.). *Phonetically driven phonology*. Cambridge University Press.

Bliese, Loren F. 1981. *A Generative Grammar of Afar.* Summer Institute of Linguistics, University of Texas at Arlington.

Blust, Robert. 1990. Three recurrent changes in Oceanic languages. In J.H.C.S. Davidson (ed.), *Pacific island languages: Essays in honour of G. B. Milner.* London: University of London, School of Oriental and African Studies. 7-28.

Blust, Robert. 2000. Chamorro historical phonology. *Oceanic Linguistics* 39: 83-122.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, DC: Georgetown University Press.

Crosswhite, Katherine. 1999. Inter-paradigmatic homophony avoidance in two dialects of Slavic. *UCLA Working Papers in Linguistics 1, Papers in Phonology 2*. 48-67.

Dell, François, and Mohamed Elmedlaoui. 1996. On consonant releases in Imdlawn Tashlhiyt Berber. *Linguistics* 34: 357-395.

Erwin, Wallace M. 1963. A Short reference grammar of Iraqi Arabic. Washington, DC: Georgetown University Press.

Fortescue, Michael, Steven Jacobson, and Lawrence Kaplan. 1994. *Comparative Eskimo Dictionary with Aleut Cognates*. Fairbanks: Alaska Native Language

Center.
Gafos, Diamandis. 1998. Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* 16: 223-278.
Garrett, Andrew and Juliette Blevins. To appear. Morphophonological analogy. In Sharon Inkelas and Kristin Hanson (eds.). *The nature of the word: Essays in Honor of Paul Kiparsky*. MIT Press.
Glinert, Lewis. 1989. *The grammar of Modern Hebrew*. Cambridge: Cambridge University Press.
Goddard, Ives. 1979. *Delaware verbal morphology: A descriptive and comparative study*. New York: Garland.
Goddard, Ives. 1982. The historical phonology of Munsee. *IJAL* 48: 16-48.
Hayward, Dick. 1984. *The Arbore language: a first investigation.* Hamburg: Helmut Buske Verlag.
Hoijer, Harry. 1933. *Tonkawa: An Indian language of Texas.* Extract from Handbook of American Indian Languages, Vol. III. New York: Columbia University Press.
Hoijer, Harry. 1946. Tonkawa - An Indian language of Texas. In C. Osgood (ed.), *Linguistic Structures of Native America.* New York: Viking Fund Publications in Anthropology 6.
Hoijer, Harry. 1949. *An analytical dictionary of the Tonkawa language.* University of California Publications in Linguistics 5(1). Berkeley: University of California Press.
Hoijer, Harry. 1972. *Tonkawa Texts.* University of California Publications in Linguistics 73. Berkeley: University of California Press.
Jacobson, Steven A. 1984. *Yup'ik Eskimo Dictionary.* Fairbanks: Alaska Native Language Center.
Kager, Rene. 1999. *Optimality Theory.* Cambridge: Cambridge University Press.
Keer, Edward W. 1999. *Geminates, the OCP and the Nature of CON.* PhD dissertation, Rutgers University, New Brunswick, NJ.
Kenstowicz, Michael. 1996. Base identity and uniform exponence: alternatives to cyclicity. In J. Durand and S. Laks (eds.), *Current Trends in Phonology: Models and Methods.* European Studies Research Institute and University of Salford: 363-393.
Kisseberth, Charles. 1970. Vowel elision in Tonkawa and derivational constraints. In J. Sadock and A Vanek (eds.), *Studies presented to Robert B. Lees by his Students.* Papers in Linguistics, Monograph Series I. Champaign, Ill: Linguistic Research, Inc.
Lin, Yen-Hwei. 1997. Syllabic and moraic structures in Piro. *Phonology* 14:403-36.
Malone, Joseph. 1986. The obligatory contour principle, tier conflation and bracket erasure: A new alignment from Hebrew. Ms., Barnard College.
Malone, Joseph. 1993. *Tiberian Hebrew phonology.* Winona Lake, IN: Eisenbrauns.
Matteson, E. 1965. *The Piro (Arawakan) language*. Berkeley: University of California Press.
McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12: 373-418.
McCarthy, John J. 1982. Prosodic templates, morphemic templates, and morphemic tiers. In H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representations I.* Dordrecht: Foris.

McCarthy, John J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17:207-63.
McCarthy, John J. 2002. *A thematic guide to optimality theory.* Cambridge: Cambridge University Press.
Miyaoka, O. 1971. On syllable modification and quantity in Yuk phonology. *IJAL* 37: 219-226.
Odden, David. 1988. Anti Antigemination and the OCP. *Linguistic Inquiry* 19: 451-75.
Reed, E. Irene, Osahito Miyaoka, Steven Jacobson, Paschal Afcan, and Michael Krauss. 1977. *Yup'ik Eskimo grammar.* Fairbanks: Alaska Native Language Center and Yup'ik Language Workshop.
Rose, Sharon. 2000. Rethinking geminates, long-distance geminates and the OCP. *Linguistic Inquiry* 31: 85-122.
Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Janet Pierrehumbert and Michael Broe (eds.), Laboratory Phonology V: Acquisition and the lexicon. Cambridge: Cambridge University Press. 313-34.
Thurneysen, R. 1980. *A grammar of Old Irish.* Dublin: The Dublin Institute for Advanced Studies.
Wise, H. 1983. Some functionally motivated rules in Tunisian phonology. Journal of Linguistics 19: 165-181.
Woodbury, Anthony C. 1982. *Study of the Chevak dialect of Yup'ik Eskimo.* Phd dissertation, University of California, Berkeley.
Woodbury, Anthony C. 1987. Meaningful phonological processes: A consideration of Central Alaskan Yupik Eskimo prosody. *Language* 63:685-740.
Yip, Moira. 1998. Identity avoidance in phonology and morphology. In Steven LaPointe, Diane Brentari, and Patrick Farrell (eds.), *Morphology and its relations to phonology and syntax.* Stanford, CA: CSLI Publications: 216-46.

Department of Linguistics
1203 Dwinelle Hall
University of California, Berkeley
Berkeley, CA 94720-2650

jblevins@uclink.berkeley.edu

Children's Unnatural Phonology

EUGENE BUCKLEY
University of Pennsylvania

Most approaches to generative phonology grant a privileged status to "natural" phonological rules, those that have a clear phonetic motivation. In Optimality Theory, for example, children are typically assumed to begin language learning with markedness constraints ranked above faithfulness constraints (see Smolensky 1996). The assumption – often not made explicit — is that phonetically arbitrary processes require some other apparatus in the grammar, perhaps the creation of a language-specific constraint, and consequently should pose a more difficult learning task for the child.

I argue that this prediction is incorrect, based on the diversity of child phonology. The common occurrence of natural patterns can be explained without reference to grammatical constraints, while the unnatural patterns that deviate from the adult model require abandonment of such constraints (cf. Hale and Reiss 1998, 2000, Blevins to appear for similar arguments).

1. Naturalness

The existence of a great many alternative explanations for children's errors makes it difficult to draw firm conclusions about the role of naturalness in the relative ease of acquisition of various contrasts. For example, if we observe that certain sounds or distinctions are relatively more difficult to perceive, there is no need to represent these sounds as "difficult" in the cognitive model of phonological knowledge: that difficulty asserts itself in the child's auditory-perceptual system, before there is even an opportunity to extract the correct patterns from the data perceived.

Consider the very common deletion of a coda consonant (cf. Branigan 1976, Fikkert 1994, Demuth 1995); these words from Vihman (1996) illustrate CV outputs in children learning English.

(1)	a.	[dɔ]	'dog'
	b.	[ʔæ]	'apple'
	c.	[du]	'juice'

The problem with arguments from such examples, however, is that the explanation could easily lie outside the mental grammar (see Hale and Reiss 1998). Here are some possible alternative explanations.

(2) a. Less robust cues for perceiving a coda consonant in the absence of a release (cf. Steriade 1999).

b. Difficulty of a voicing gesture in final position articulating a syllable-closing gesture relative to a released consonant (for voicing, see Westbury and Keating 1986).
c. Difficulty coordinating or processing two consonantal gestures in one syllable or word, leading to the sacrifice of the less salient one (cf. Vihman 1978, Berg 1992).
d. A quantal effect between the maximally distinct categories of vowel and consonant (Beckman and Edwards 2000), leading to an "oscillation frame" of C's and V's (McNeilage and Davis 1990, 1999).

To the extent that strong tendencies in early child language, such as a preference for open syllables, can be attributed to non-grammatical factors, no grammatical conclusion can be drawn. An additional problem for markedness is that highly marked processes are also found in child speech.

2. Onset / Coda (A)symmetries

Despite the strength of the tendency toward CV syllables, "[t]here is no reason to believe that all children initially require all syllables to have onsets" (Bernhardt and Stemberger 1998: 371). In fact, there are many situations in which onsetless syllables occur in child language, including of course languages with onsetless adult forms. For Inês (0;11–1;5) learning Portuguese (Freitas 1996), even target onsets are often omitted or a vowel without an onset is inserted; matching results are reported for two other children.

(3)		Child	Adult		
	a.	[ˈpo]	[ˈpõj̃]	*põe*	'put'
	b.	[ɑˈti]	[ɑˈki]	*aqui*	'here'
	c.	[ˈa:]	[ˈagwɑ]	*água*	'water'
	d.	[ɑˈpɑ]	[ˈpãw̃]	*pão*	'bread'
	e.	[eˈɑ]	[ʒuˈãw̃]	*João*	(name)

Similarly, Amahl learning English (Smith 1973) often deleted [h] and sibilants during his third year, leading to onsetless syllables, many with codas.

(4)	a.	[ɛn]	'hand'	[ʌn]	'sun'
	b.	[ɛt]	'head'	[a:p]	'sharp'
	c.	[ʌ:t]	'hurt'	[up]	'soup, soap'
	d.	[ɛŋu]	'handle'	[ʊgə]	'sugar'

In this case a competing constraint against the offending segments is responsible. (There might also be glottal stop onsets, untranscribed.) But other examples show a more positive attention to codas.

Fey and Gandour (1980) present an instance of a more complex output innovated by the child. Observed from 1;9–2;2, Lasan produced nearly all word-final voiced stop targets with a following syllabic nasal. (A small number of such targets were deleted or modified to other consonants.)

(5)	a.	[dabm̩]	'stub'
	b.	[dædn̩]	'dad'
	c.	[bəgŋ̍]	'bug'

The authors (p. 77) observe that, in general for Lasan, "word-final position had special status for marking perceived phonological distinctions" — for example, final voiceless stops were consistently realized as aspirated or released stops with the correct place of articulation. The nasal release of the voiced stop was the only means available to Lasan at the time to preserve the voicing contrast: he had apparently not yet mastered the articulatory maneuver of voicing a word-final stop. Of course, the frequency of adult languages that neutralize this distinction attests to the relative difficulty of maintaining the contrast in final position. Fey and Gandour also note that, while the functional and phonetic motivation of the rule is clear, the rarity of analogous processes in adult languages casts doubt on the reasonableness of treating this as a "natural" process (in the sense of Stampe 1979) or "unmarked" more generally (as in Optimality Theory: Prince and Smolensky 1993).

Similarly, Menn (1971) shows that Daniel (from 22.5 to 24 months) deleted fricatives in the onset but permitted them in the coda.

(6)	a.	[it]	'seat'	[dos]	'toast'
	b.	[iz, is]	'cheese'	[æʃ]	'watch'
	c.	[ɨʃ]	'fish'	[ejndʒ]	'change'
	d.	[uz]	'shoes'	[ufs]	'juice'

More generally, at this stage Daniel had a much larger inventory of sounds in the coda than in the onset. (Bernhardt and Stemberger 1998: 436 identify two similar cases of asymmetry.)

(7)	a.	onset inventory	[b, d, g, k, m, n]
	b.	coda inventory	[b, d, g, p, t, k, r, s, ʃ, z, f, v, dʒ, m, n, ŋ]

Further, for a syllable ending in a stop, the onset contrasts are even more severely restricted.

(8)		Adult onset	Daniel's onset	Examples
	a.	[n], [sn]	[n]	[næk] 'napkin' [næt] 'snap'
	b.	[m], stop, affricate	homorganic stop	[dit] 'meat' [gʊk] 'book'
	c.	fricative, approximant	zero	[iv] 'Steve' [ajd] 'ride'

This situation violates the adult-language observation that onsets permit more contrasts than codas (though there are exceptions even in adult phonologies: Blevins to appear). As noted by Branigan (1976) and Menn (1978), some children simply seem to concentrate more on the beginnings or ends of words. Such variation supports a more cognitive approach to phonological development, in which the child is a creative problem solver (cf. Macken and Ferguson 1983, Menyuk, Menn, and Silber 1986), over a more deterministic universal grammar.

Inkelas and Rose (to appear) present a persuasive argument for the effect of a child's physiological limitations on the creation of unexpected phonological rules. Specifically, "velar fronting" replaces alveolars with velars in the onsets of stressed syllables, exactly where we expect the greatest faithfulness, while coda and unstressed velars are produced accurately.

(9)	a.	[tʰʌp]	'cup'
	b.	[do:]	'go'
	c.	[tʰʊk]	'cook'
	d.	[ˈtokoˌnʌt]	'coconut'

Their explanation for this pattern, which is unattested in adult languages and thus quite unnatural, is that the child recognizes the greater amplitude of consonant gestures in the onset of an adult stressed syllable; but the relatively large tongue size in a child below two years of age makes a large-amplitude dorsal articulation difficult. The end result of the child's attention to this phonotactic distribution is a change to an alveolar articulation in exactly the circumstance where the greater gesture is needed. The rule will normally disappear as the child's mouth achieves more adult-like proportions; but the fact that it occurs at all shows that the phonology has to accommodate such "unnatural" processes.

Berg (1992) discusses labial harmony by Melanie (2;7–2;11), learning German. The general direction of this harmony is from right to left, which includes onset features sacrificed to those of a coda.

(10)	a.	[po:mas]	[to:mas] *Thomas*	(name)
	b.	[bi:bən]	[ʃi:bən] *schieben*	'to push'
	c.	[me:mən]	[ne:mən] *nehmen*	'to take'
	d.	[bɛlp]	[gɛlp] *Gelb*	'yellow'
	e.	[pɔmt]	[kɔmt] *kommt*	'comes'
	f.	[bo:m]	[do:m] *Dom*	'cathedral'

The direction of assimilation indicates that "the most difficult position for her is the word onset" (p. 241); again, this is a case of onset constrasts being reduced relative to coda (as well as intervocalic) contrasts. Consonant harmony is a classic example of child phonology differing from adult patterns, but the directionality (which is the same for most other children; cf. Vihman 1978) also contradicts the usual primacy of onsets and their contrasts. Even if the salience of intervocalic vs. word-initial contrasts is the original basis for the directionality, what's noteworthy is that the resulting phonological rule is generalized to apply from coda to onset as well, in defiance of expectations based on naturalness.

Stemberger (1996) and Bernhardt and Stemberger (1998: 383) report that Morgan optionally inserted glottal stop after a word-final vowel until 1;4, thereby creating codas.

(11)	a.	[mama] ~ [mamaʔ]	'mama'
	b.	[lɪlɪ] ~ [lɪlɪʔ]	'lizard'
	c.	[mi:] ~ [miʔ]	'me'

That she was particularly interested in word-endings is confirmed by the fact that "from the first word on, Morgan never deleted word-final codas."

The requirement for a final coda finds a parallel in certain languages, such as glottal-final words in Makassarese (Aronoff et al. 1987, McCarthy and Prince

1993) and C-final stems in Arabic (McCarthy and Prince 1990), but such patterns are hardly "natural" in the usual sense. In Makassarese, a preference for a final consonant, specifically [ʔ], is normally inhibited by faithfulness to the input, but arises in conjunction with other processes: the insertion of a final epenthetic vowel and partial reduplication. The result is strikingly similar to Morgan's rule.

(12)	a.	/rantas/	rántasaʔ	'dirty'
	b.	/jamal/	jámalaʔ	'naughty'
	c.	/manara/	manaʔ - manára	'sort of tower'
	d.	/balao/	balaʔ - baláo	'toy rat'

Note also the insertion of glottal stop after an utterance-final short vowel in Japanese (Vance 1987: 12). What these processes show — entirely consistent with the claims of this paper — is that the set of phonological tools available to adult languages is also available to children, and can be used to construct both natural and unnatural distributions of sounds.

Stern and Stern (1907) describe productions of their German-learning son Günther (beginning 2;1–2;2 and lasting about 6 months) with a strong tendency to replace word-initial consonants with [h] or zero. Many of the replaced single-consonant onsets — including [p, t, k, b, g, f, v, z, r, l, m] — are successfully produced in other words, but the majority of lexical items show deletion or replacement.

(13)	a.	[ˈhas] *haß*	[ˈnas] *naß*	'wet'
	b.	[ˈhɛlt] *held*	[ˈgɛlt] *Geld*	'money'
	c.	[ˈhu:] *hū*	[ˈʃu:] *Schuh*	'shoe'
	d.	[ˈhʊmpf] *humpf*	[ˈʃtrʊmpf] *Strumpf*	'stocking'
	e.	[əˈhɔfən] *ehoffen*	[gəˈtrɔfən] *getroffen*	'met, struck'
	f.	[ɛˈhanda] *ähanda*	[vɛˈranda] *Veranda*	'veranda'
	g.	[aˈe:t] *aēt*	[paˈke:t] *Packet [Paket]*	'packet'
	h.	[iˈi:çən] *i-ichen*	[maˈri:çən] *Mariechen*	'Marie (dim.)'
	i.	[uˈɛla] *uëlla*	[kuˈrɛla] *Kurella*	(name)
	j.	[ˈu:l] *ūl*	[ˈʃtu:l] *Stuhl*	'chair'
	k.	[ˈɛks] *ex*	[ˈdaks] *Dachs*	'badger'

The authors give only orthographic representations, shown here in italics; the phonetic transcriptions are my best estimate of the actual spoken form, though some details are uncertain. Some words are transcribed with initial <ch>, which the authors (p. 95) intend as the velar fricative [x].

(14)	a.	[ˈha:tɐ, ˈxa:tɐ] *hātă, chātă*	[ˈfa:tɐ] *Vater*	'father'
	b.	[ˈho:s, ˈxo:s] *hōß, chōß*	[ˈgro:s] *groß*	'big'
	c.	[ˈxe:ɐ] *chēr*	[ˈme:ɐ] *mehr*	'more'
	d.	[ˈxʏntɐ] *chünther*	[ˈgʏntɐ] *Günther*	(name)
	e.	[ˈaʊˌxuɐç] *auchurch*	[ˈmaʊlˌvuɐf] *Maulwurf*	'mole'

Truncations of longer words with non-initial stress sometimes retain an onset (cf. 14a-b) but often have a simplified onset despite an available consonant.

(15)	a.	[ˈpʊt] *put*	[kaˈpʊt] *kaput [kaputt]*	'broken'
	b.	[ˈla:də] *lade*	[ʃokoˈla:də] *Schokolade*	'chocolate'

c.	[ˈhant] *hant*	[eleˈfant] *Elefant*	‘elephant’
d.	[ˈa:t] *āt*	[zaˈla:t] *Salat*	‘salad’
e.	[ˈo:tə] *ōte*	[koˈmo:də] *Kommode*	‘dresser’
f.	[ˈy:tlɪç] *ütlich*	[gəˈmy:tlɪç] *gemütlich*	‘cozy’

Words that already conform to the pattern tend to be reproduced accurately.

(16)	a. [ˈhu:t] *hut*	[ˈhu:t] *Hut*	‘hat’
	b. [ˈhals] *hals*	[ˈhals] *Hals*	‘neck, throat’
	c. [ˈɔks] *ochs*	[ˈɔks] *Ochs(e)*	‘ox’
	d. [ˈi:gəl] *igel*	[ˈi:gəl] *Igel*	‘hedgehog’

It may be that the apparently vowel-initial words begin with a glottal stop (as found in adult pronuncation); the “preferred” onsets would then be the related group [ʔ, h, x]. In a few cases, one of these onsets is substituted for another.

(17)	a. [ˈxaɪs] *cheiß*	[ˈhaɪs] *heiß*	‘hot’
	b. [ˈha:lɐ] *hāler*	[ˈa:dlɐ] *Adler*	‘eagle’

What remains certain is that the consonant constrasts are reduced in the onset rather than in the coda, very much contrary to the norm. Günther seems to prefer complete non-identity rather than partial identity, which is interesting in itself. But what matters most in the present context is that this restriction applies to onsets, not to codas, which he usually produces faithfully, even when they contain clusters.

The possibility of a bias against onsets in some child languages indicates that the human phonological capacity does not rule out such grammars, even if (for extragrammatical reasons) they do not arise frequently. By the same token, we expect adult languages to show a similar bias, if only rarely. And indeed, there have been claims that some languages, such as Barra Gaelic (Clements 1986) and Oykangand (Sommer 1970), prefer VC syllables sometimes or always. These are the basic syllable types of Oykangand, following Sommer.

(18)	a.	VC	/ef/	‘tongue’
	b.	VCC	/er̃k/	‘ground, place’
	c.	VCCC	/algŋ/	‘tooth’
	d.	VCCCC	/albmb/	‘opossum’

While the various pressures favoring CV syllables ensure that such coda-favoring systems are unlikely to arise in language change, under the right circumstances this computationally possible, albeit typologically rare pattern will emerge (Blevins to appear). The Oykangand pattern arose due to loss of initial consonants (itself rather similar to Günther’s rule). The interest of child language is that the creation of a new grammar based on impoverished data liberates the computational capacity of the grammar from the typological filtering of slow change over time, affording a unique look at the underused corners of the phonological toolkit.

3. Innovative Unnaturalness

It has been argued that the unnaturalness of an adult alternation does not impede the learning of a process by children; what matters is whether the adult rule is regular and well attested in the data (see Buckley 2002 and references therein).

Unnatural patterns that are innovated by children (rather than being faithful imitations of unnatural adult patterns) contribute to the same conclusion. Beckman and Edwards (2000: 242) state that

> even after the child has mastered the articulation of a specific phoneme in one set of words, generalization of the motor pattern to other forms containing the phoneme depends on the frequency with which the phone has been encountered in the immediate phonetic context exemplified by the novel form.

What this means is that the child's (apparent) mastery of particular articulations does not give us a clear view of innate language ability, but rather is significantly affected by the vocabulary of a language, and is therefore very much affected by centuries of accreted changes and historical coincidences.

Priestly (1977) describes the classic example of a [CVjVC] template for Christopher (1;10–2;1); this is certainly not natural or typical for children, but it resembles some adult templates and clearly is within the scope of the phonological capacity. The choice of which consonant occupies the final slot is apparently driven by its saliency or familiarity (p. 60).

(19)	a.	[fajan]	'flannel'	[sɛjan]	'whale'
	b.	[hajaŋ]	'hanger'	[pijat]	'peanut'
	c.	[bajak]	'blanket'	[dajak, dajan]	'dragon'

A full 25% of "ordinary" disyllabic forms reported by Priestly (in a sample from week 4 of the study) contain a medial [j], and these frequent inputs likely served as the model for the template.

(20)	a.	[lǻjən]	'lion'
	b.	[wɛ̃́jəl]	'whale'
	c.	[fǻjə]	'fire'

Whatever the inspiration, Christopher's forms resemble results from left-to-right and edge-in association in adult templatic morphologies (cf. Yip 1988). The presence of a fixed element in a template also attested, as with certain verbs of Classical Arabic (Wehr 1971). An example is the /w/ in Form XII of the verb.

(21)	a.	ħadib-	'be convex'	iħda<u>w</u>dab-	'be crooked, vaulted'
	b.	ɣariq-	'plunge, be immersed'	iɣra<u>w</u>raq-	'be bathed (in tears)'

Adult grammars such as Arabic show the computational possibility of such templatic patterns, which are available to be exploited by children given the right skewing of the input, or even some predisposition on the part of a child.

Stemberger (1992: 178) and Bernhardt and Stemberger (1998: 403) report that Gwendolyn went through a stage at 3;3 in which she created novel clusters. In earlier pronunciations, sonorants following a stop were deleted. Later, the sonorants /r, l, w/ were variably restored as [w], not only to words that have a cluster in the adult form but in fact to any stop-initial word.

(22)			Deletion	Overgeneralization
	a.	'tree'	[t^hi:]	[t^hi:] ~ [t^hwi:]
	b.	'top', 'stop'	[t^ha:p]	[t^ha:p] ~ [t^hwa:p]

c. 'book' [bət] [bət] ~ [bwət]

These novel pronunciations have nothing to do with markedness, but rather a variable hypercorrection of forms with a stop in the onset. Adult sound change can, of course, lead to a more complex syllable structure; a simple case is vowel devoicing or deletion in the word *potato* when pronounced with initial [p^ht^h]. In the adult and child changes, we see that a particular change can proceed without regard to supposed markedness pressures that would oppose it.

A different sort of evidence of how phonological markedness is subordinated to other aspects of the input comes from Labov's (1989) work on the acquisition of sociolinguistic variables. He studied children aged 4–9, learning the English variable processes *-t,d* deletion and g-dropping in *-ing*. Among his findings was that the children "first show the social and stylistic constraints on variation, then the language-specific grammatical and articulatory constraints" (p. 96). One articulatory constraint is the effect of adjacent consonants as defined by relative sonority — quite a natural pattern, yet learned later than the arbitrary social factors governing the variation.

Finally, in a particularly interesting example, Bernhardt and Stemberger (1998: 639) report that Morgan misanalyzed the palatalization pattern found in English phrases such as *need you*, where the alveolar stop preceding a form of *you(r)* becomes palato-alveolar. The clear analysis of the adult alternation (*pace* Bernhardt and Stemberger) is an assimilation in place of articulation, with consequent affrication for palatal stops in English. This, of course, is an extremely common and "natural" process.

Morgan's output shows that (at an unspecified age, lasting for about 2 months) she saw it instead as an assimilation in obstruency: The initial /y/ of the pronominal becomes a palato-alveolar fricative that agrees with a preceding obstruent in voicing, but is unchanged after a sonorant.

(23)	a.	need [ʒ]ou	keep [ʃ]ou	spin [y]ou
	b.	hug [ʒ]ou	want [ʃ]ou	comb [y]ou
	c.	love [ʒ]ou	like [ʃ]ou	

Presumably this reinterpretation was affected by inflected forms in adult speech that lead to a similar outcome, though only in the presence of the sibilant suffix.

(24)	a.	[lʌvʒu]	'loves you'
	b.	[lʌvyu]	'love you'

What matters is that naturalness seems to have played no role in the new rule; instead, she seized upon a more general formulation of the process, affecting the major class feature [sonorant] rather than a specific detail of place of articulation.

Assimilation of [–sonorant] is quite poorly motivated in adult phonologies, so much so that it is typically rendered impossible by standard feature geometries (cf. Schein and Steriade 1986, McCarthy 1988, Clements and Hume 1995). Palatal assimilation, on the other hand, is one of the most common processes in the world's languages (Bhat 1978). Yet Morgan actually formed the wrong generalization, choosing a much less natural rule for the process she observed.

From the point of view of phonological theory, this is significant: Such a process does appear to be possible, even if rarely found in adult languages. This is all the more reason to focus on phonology as the characterization of possible sound inventories and processes, rather than common or "natural" processes whose frequency results from their likelihood of arising by historical change (Hale and Reiss 1998, 2000, Blevins to appear).

4. Experimental Evidence

Various attempts have been made to determine experimentally what kinds of phonological generalizations are learned more easily than others. While the matter is a difficult one to test, particularly with regard to children first acquiring a language, there is support for the idea that unnatural processes are learned well. Here I discuss just two representative studies.

Jusczyk, Smolensky, and Allocco (2002) have claimed that infants are predisposed to prefer unmarked patterns, due to a universal grammar that encodes markedness. The natural rule under consideration was nasal place assimilation. Subjects of various ages were exposed to stimulus sets consisting of three items: two isolation forms followed by their concatenation, with or without a phonological change.

(25)	unmarked:	[on, pa, ompa]	(assimilation)
	marked:	[un, ber, unber]	(no assimilation)

The authors found longer listening for the assimilated (natural) triads than for the unnatural ones; they concluded that infants have a predisposition to prefer such unmarked relations. One problem is that this listing approach has not previously been used in infant studies, and it is far from clear that the sequences will be interpreted as rule inputs and outputs, or as related forms.

Setting aside this methodological objection, the vocabulary of English has many examples of assimilated sequences such as [mp], which far outnumber unassimilated [nb]; thus there could be an influence from the ambient language quite independent of any innate preference. The authors note, however, that is debatable whether infants as young as some of their subjects (4 months) are sensitive to native language patterns.

Even discounting the ambient language, there is another explanation independent of markedness: The greater representational simplicity of the assimilated form (i.e. a single intervocalic place of articulation) could also be at work. If we assume that these infants have adult-like sound representations at all, then we should assume that some may be simpler than others, which they might prefer as "better" (if not "more natural"). An innate preference for natural rules is not necessary as an explanation, even if we accept other aspects of the experimental design.

Chambers, Onishi, and Fisher (2003) tests whether 16.5-month-old infants could learn arbitrary generalizations about phonotactics. Two sets of consonants with no common features were arranged in CVC syllables.

(26)	group 1	[b, k, m, t, f]
	group 2	[p, g, n, tʃ, s]

The choice of vowel [ɪ, æ] correlated with the order of consonants, differently in two lists. Infants were familiarized with one of the two, then exposed to randomized stimuli representing both patterns.

(27) list 1 [bɪp, kɪg, pæb, gæk] etc.
list 2 [bæp, kæg, pɪb, gɪk] etc.

The expectation is that they will listen longer to unfamiliar stimuli (Saffran et al. 1996, Marcus et al. 1999); relative familiarity depends on having learned the pattern in the familiarization set. The results show that the infants had internalized the arbitrary phonotactics, and listened longer to the novel forms despite the fact that there was no natural phonological basis for the patterns. Since infants can quickly absorb arbitrary phonotactic patterns after rather brief exposure; it should come as no surprise that a child learning a language could also learn an unnatural pattern in the data.

While we must await more data before drawing firm conclusions, these results suggest that infants can learn natural and unnatural patterns after brief exposure, supporting the other evidence presented here that naturalness is not a significant factor in the learning of phonological rules.

References

Aronoff, Mark, Azhar Arsyad, Hasan Basri, and Ellen Broselow. 1987. Tier configuration in Makassarese reduplication. *Chicago Linguistic Society 23, Papers from the Parasession on Autosegmental and Metrical Phonology*, 1-15.

Beckman, Mary E., and Jan Edwards. 2000. The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development* 71: 240-249.

Berg, Thomas. 1992. Phonological harmony as a processing problem. *Journal of Child Language* 19: 225-257.

Bernhardt, Barbara Handford, and Joseph P. Stemberger. 1998. *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. San Diego, Calif.: Academic Press.

Bhat, D.N.S 1978 A general study of palatalization. *Universals of human language, vol. 2: Phonology*, 47-92. Stanford University Press.

Blevins, Juliette. To appear. *Evolutionary phonology: the emergence of sound patterns*. Cambridge University Press.

Branigan, George. 1976. Syllabic structure and the acquisition of consonants: The great conspiracy in word formation. *Journal of Psycholinguistic Research* 5: 117-133.

Buckley, Eugene. 2002. Rule naturalness and the acquisition of phonology. Second North American Phonology Conference, Concordia University, Montréal, April 25-28.

Chambers, Kyle E., Kristine H. Onishi, and Cynthia Fisher. 2003. Infants learn phonotactic regularities from brief auditory experience. *Cognition* 87: B69-B77.

Clements, G.N. 1986. Syllabification and epenthesis in the Barra dialect of Gaelic. In Koen Bogers, Harry van der Hulst, and Maarten Mous (eds.) *The phonological representation of suprasegmentals*, 317-336. Dordrecht: Foris.

Clements, G.N., and Elizabeth V. Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.) *The handbook of phonological theory*, 245-306. Cambridge, Mass.: Blackwell.

Demuth, Katherine. 1995. Markedness and the development of prosodic structure. In Jill Beckman (ed.) *Proceedings of the North East Linguistic Society* 25, 13-25. Amherst, MA: GLSA, University of Massachusetts.
Fey, Marc E., and Jack Gandour. 1982. Rule discovery in phonological acquisition. *Journal of Child Language* 9: 71-81.
Fikkert, Paula. 1994. *On the acquisition of prosodic structure*. Leiden: HIL.
Freitas, M. João. 1996. Onsets in early productions. In Barbara Bernhardt, John Gilbert, and David Ingram (eds.) *Proceedings of the UBC International Conference on Phonological Acquisition*, 76-84. Somerville, Mass.: Cascadilla Press.
Hale, Mark, and Charles Reiss. 1998. Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29: 656-683.
Hale, Mark, and Charles Reiss. 2000. Phonology as cognition. In Noel Burton-Roberts, Philip Carr, and Gerard Docherty (eds.) *Phonological knowledge: Conceptual and empirical issues*, 161-184. Oxford University Press.
Inkelas, Sharon, and Yvan Rose, to appear. *Proceedings of the Boston University Child Language Development Conference* 27.
Jusczyk, Peter, Paul Smolensky, and Theresa Allocco. 2002. How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition* 10: 31-73.
Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1: 85-97.
Macken, Marlys A. and Charles A. Ferguson. 1983. Cognitive aspects of phonological development: Model, evidence and issues. In Keith E. Nelson (ed.) *Children's language* 4, 255-282. Hillsdale, N.J.: Lawrence Erlbaum.
MacWhinney, Brian. 1978. *The acquisition of morphophonology*. [Monographs of the Society for Research in Child Development, serial no. 174, vol. 43, nos. 1-2.] University of Chicago Press.
Marcus, G.F., S. Vijayan, S. Bandli Rao, and P.M. Vishton. 1999. Rule learning by seven-month-old infants. *Science* 283: 77-80.
McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 45: 84-108.
McCarthy, John J., and Alan S. Prince. 1990. Prosodic morphology and templatic morphology. In Mushira Eid and John McCarthy (eds.) *Proceedings of the Second Annual Symposium on Arabic Linguistics*, 1-54. Amsterdam: John Benjamins.
McCarthy, John J., and Alan S. Prince. 1993. Generalized alignment. In G. Booij and J. van Marle (eds.) *Yearbook of Morphology 1993*, 79-153. Dordrecht: Kluwer.
McNeilage, P.F., and B.L. Davis. 1990. Acquisition of speech production: Frames, then content. In M. Jeannerod (ed.) *Attention and performance: Vol. 13, Motor representation and control*, 453-475. Hillsdale, N.J.: Erlbaum.
McNeilage, P.F. and B.L. Davis. 1999. Consonant (vowel) consonant sequences in early words. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, CA, 2489-2492.
Menn, Lise. 1971. Phonotactic rules in beginning speech. *Lingua* 26: 225-251.
Menn, Lise. 1978. Phonological units in beginning speech. In A. Bell and J.B. Hooper (eds.) *Syllables and Segments*, 315-334. Amsterdam: North Holland.
Menyuk, Paula, Lise Menn, and Ronnie Silber. 1986. Early strategies for the perception and production of words and sounds. In Paul Fletcher and Michael

Garman (eds.) *Language acquisition: Studies in first language development*, 198-222. Cambridge University Press.
Priestly, T.M.S. 1977. One idiosyncratic strategy in the acquisition of phonology. *Journal of Child Language* 4: 45-66.
Prince, Alan S., and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado, Boulder.
Saffran, J., R. Aslin, and E. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274: 1926.
Schein, Barry, and Donca Steriade. 1986. On geminates. *Linguistic Inquiry* 17: 691-744.
Smith, N.V. 1973. *The acquisition of phonology: A case study*. Cambridge University Press.
Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27: 720-731.
Sommer, Bruce. 1970. An Australian language without CV syllables. *International Journal of American Linguistics* 36: 57-58.
Stampe, David. 1979. *A dissertation on natural phonology*. New York: Garland.
Stemberger, Joseph P. 1992. A connectionist view of child phonology: Phonological processing without phonological processes. In Charles A. Ferguson, Lise Menn, and Carol Stoel-Gammon (eds.) *Phonological development: Models, research, implications*, 165-189. Timonium, Md.: York Press.
Stemberger, Joseph P. 1996. Syllable structure in English, with emphasis on codas. In Barbara Bernhardt, John Gilbert, and David Ingram (eds.) *Proceedings of the UBC International Conference on Phonological Acquisition*, 62-75. Somerville, Mass.: Cascadilla Press.
Steriade, Donca. 1999. Phonetics in phonology: The case of laryngeal neutralization. In M.K. Gordon (ed.) *UCLA Working Papers in Linguistics* 2: 25-146.
Stern, Clara, and William Stern. 1907. *Die Kindersprache: Eine psychologische und sprachtheoretische Untersuchung*. Leipzig: Johann Ambrosius Barth.
Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany: State University of New York Press.
Vihman, Marilyn May. 1978. Consonant harmony: Its scope and function in child language. In J.H. Greenberg (ed.) *Universals of Human Language, Vol. 2: Phonology*, 281-334. Stanford University Press.
Vihman, Marilyn May. 1996. *Phonological development: The origins of language in the child*. Cambridge, Mass.: Blackwell.
Wehr, Hans. 1971. *A dictionary of Modern Written Arabic*. Third edition, ed. J. Milton Cowan. Wiesbaden: Otto Harrassowitz.
Westbury, John R., and Patricia A. Keating. 1986. On the Naturalness of Stop Consonant Voicing. *Journal of Linguistics* 22: 145-166.
Yip, Moira. 1988. Template morphology and the direction of association. *Natural Language and Linguistic Theory* 6: 551-577.

Eugene Buckley
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104-6305

gene@unagi.cis.upenn.edu

Consonant Confusability: An MEG Study*

VALENTINE HACQUARD and MARY ANN WALTER
Massachusetts Institute of Technology

0. Introduction

Numerous behavioral experiments have investigated consonant confusability by evaluating error rates in identification of phonemes masked with noise. These studies have found that the pair of nasals, /m/ and /n/, is more confusable than the corresponding oral pair /b/ and /d/ (Miller & Nicely 1955, Wang & Bilger 1973). The members of the two pairs differ in precisely the same way in terms of abstract phonological representations – both involve a change from labial to coronal place of articulation. Thus the difference in confusability between the two pairs of consonants has been explained in acoustic terms. The nasal pair is acoustically more similar for at least two reasons. First, there is a perseveration or coarticulation effect of the nasal on the following vowel. Second, there is a reduction of acoustic distance (mainly with respect to the second formant) between nasalized vowel transitions compared to oral vowel transitions (Wright 1986). Due to this acoustic similarity, the nasal pair is more confusable and, we assume, perceived to be more similar.

In this study, we examine subjects' perception of similarity for the same group of contrasts – /ma/ and /na/ versus /ba/ and /da/ – but with a much finer-grained time-course and level of discrimination than a behavioral judgment. We use magnetoencephalography (MEG) brain-imaging to evaluate variation in two early auditorily-evoked neural components – the M100 response and the mismatch field response (MMF). MEG records the magnetic field generated by electrical activity in the brain, and is able to provide a millisecond-by-millisecond picture of brain activity. Our goal is to shed light on the earliest available stage of perception with respect to the relative roles of abstract phonological features and acoustic similarity. If the latter proves relevant at this stage, then brain-imaging data can be used to test proposals for similarity metrics, and answer questions about the basic units of contrast and their relation to each other.

---

* We would like to thank Alec Marantz, Diana Sonnenreich, Donca Steriade, Karen Froud, Linnaea Stockall, and Pranav Anand. We are extremely indebted to Ken Stevens for help synthesizing the stimuli.

1. Two models of similarity

1.1. An acoustic-based model

Much current work in phonology motivates phonological processes by appealing to perceptual, acoustic factors. Similarity along acoustic dimensions has been argued to underlie phonological processes such as metathesis, assimilation, and phonotactic constraints. Hume (1998) argues that confusability considerations motivate metathesis processes, so that a segment is positioned in such a way as to maximally enhance its perceptual salience. Steriade (1999) proposes that distance along a perceptually-based similarity metric (the P-map) correlates with assimilation processes. In the context of the confusability studies reported above, the findings of Mohanan (1993) and Jun (1995) that nasals are relatively more susceptible to place assimilation than stops generally is unsurprising.

These confusability and assimilation facts, as well as studies of offline similarity judgments (Hura et al. 1992), support a model of similarity based on acoustic, perceptual factors. Such a model is dependent on phonetic context, but independent of the phonological inventory. According to an acoustic-based model, nasal confusability is explained in perceptual terms: although nasality itself is highly salient, persevering nasality alters the F2 transition into the following vowel, which is an important cue for place of articulation.

1.2. A model based on natural classes (Frisch 1996)

Frisch et al. (1997) and Frisch (1996) propose an alternative model of similarity, which differs from the acoustic-based model in that it is independent of phonetic context, but dependent on the phonological inventory and its structure. This similarity metric is computed according to natural classes, as illustrated in (1):

(1) $$\text{Similarity} = \frac{\text{shared natural classes}}{\text{shared + unshared natural classes}}$$

The similarity measure in (1) computed for English consonants gives the values in (2), where the ranking of similarity contradicts that of the acoustic-based model: nasals are equally (if a .01 difference is insignificant) or less similar to each other than orals.

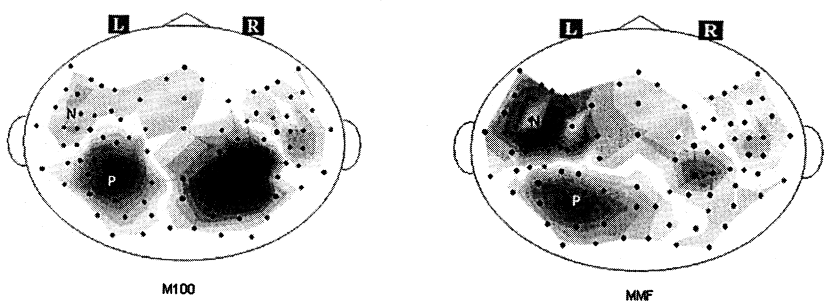
(2) b/d (= .29) > m/n (= .28), where 1 is identity (maximal similarity).

In our study, we will thus examine whether a difference in similarity, along any of the metrics discussed above, correlates with a difference in an auditory brain response. If it does, then brain data can be used to substantiate proposals for similarity metrics, such as the two proposed above, and better understand their internal organization.

2. Magnetoencaphalography (MEG)

MEG measures the magnetic fields generated by electrical activity in the brain, specifically by potentials in the apical dendrites of pyramidal cells in the cortex. The main component we look at here is the mismatch field response or MMF, an automatic, auditory brain response evoked by a deviant stimulus following a sequence of standards, peaking ~180-250 ms after the difference point between a standard and a deviant. We also look at the M100, an automatic auditorily-evoked response that peaks ~100 ms post-stimulus onset.

Figure 1 – Magnetic field distributions of the M100 and MMF in response to speech stimuli for one representative subject. The letter P indicates the positive field (emerging from the brain) and the letter N the negative field (entering the brain).



Previous MEG studies have identified several properties of the MMF. Sharma & Dorman (1999) and Phillips et al. (2000) show that the same VOT span crossing a phonemic category boundary evokes a far greater MMF than one that doesn't. Näätänen et al. (1997) show that a small acoustic difference crossing a phonemic category boundary evokes a far greater MMF than a large one that does not cross such a boundary. It appears, then, that phonological difference outweighs acoustic difference for the MMF response.

However, we would like to know whether similarity distinctions matter when category is kept constant. In our experiment, both the oral pair and the nasal pair involve crossing a phonological category (i.e., place of articulation from labial to coronal and vice-versa). If the gap within one pair differs from the gap of the other pair, then some other factor plays a role. We hypothesize that the factor is perceptual similarity.

3. Materials and methods

3.1. Participants

Seventeen subjects, all students and employees at MIT, all of whom gave informed consent, participated in the experiment. The sample included 9 males and 8 females. All subjects were right-handed and had no history of hearing or neurological disorders. Eleven were native speakers of English. Six were native speakers of languages which also have the relevant contrasts (Czech (2), Spanish, Brazilian Portuguese, Russian, and Persian). They were also fluent in English. One subject was later excluded due to technical problems that arose during the experiment.

3.2. Stimuli

Stimuli for the present experiment consisted of four CV syllables – /ba/, /da/, /ma/, /na/. The syllables were synthesized using the program KLSyn. The stimuli were correctly identified and discriminated by speakers of languages with the relevant contrasts, thus ensuring that they were good exemplars.

3.3. Procedure

Subjects lay supine in a magnetically shielded room while stimuli were presented binaurally over earphones. Evoked magnetic fields were recorded using the MIT/KIT whole-head biomagnetometer array, with 93 axial gradiometers.

Before beginning the experiment, subjects listened to a 1kHz tone presented 100 times, for the purpose of identifying and localizing the M100, and helping to select RMS sensors for the analysis (see below). We used a mismatch detection task (oddball paradigm) where a series of four identical precursor stimuli, separated by inter-stimulus intervals of 400 ms, was presented followed by a fifth stimulus, either identical to the previous four (control condition or *standard*), or different from it with respect to place of articulation of the initial consonant only (*deviant* condition). The stimuli were arranged into the eight conditions illustrated in Figure 2. The subjects heard each condition 30 times, in random order.

Trial presentation was randomized by Psyscope script. Subjects were instructed to press one button when the fifth item of each set was the same as the previous ones, and a second button if it was different. The trials were divided into six blocks of forty trials. Between blocks, the subject was given a break of self-determined duration.

Data were sampled at 1000 Hz, with acquisition between 1 and 200 Hz. The recording for each participant lasted approximately 25 minutes. The raw data was then subjected to a noise reduction routine to eliminate measured magnetic activity from external sources. Responses to stimuli were averaged by stimulus condition separately, in 700 ms windows keyed to the onset of the stimulus: 100 ms pre-, 600 ms post onset. The averaged signal was subjected to a low-pass filter below 30 Hz and adjusted to baseline using a 100 ms pre-stimulus interval.

Figure 2 – Conditions (8x30=240 trials)

1) ba	(400 ms)	ba	(400 ms)	ba	(400 ms)	ba	(400 ms)	**da**	deviant
2) da	(400 ms)	da	(400 ms)	da	(400 ms)	da	(400 ms)	da	standard
3) da	(400 ms)	da	(400 ms)	da	(400 ms)	da	(400 ms)	**ba**	deviant
4) ba	(400 ms)	ba	(400 ms)	ba	(400 ms)	ba	(400 ms)	ba	standard
5) ma	(400 ms)	ma	(400 ms)	ma	(400 ms)	ma	(400 ms)	**na**	deviant
6) na	(400 ms)	na	(400 ms)	na	(400 ms)	na	(400 ms)	na	standard
7) na	(400 ms)	na	(400 ms)	na	(400 ms)	na	(400 ms)	**ma**	deviant
8) ma	(400 ms)	ma	(400 ms)	ma	(400 ms)	ma	(400 ms)	ma	standard

3.4. Data analysis

Both button-press responses and reaction times were recorded. Reaction times were calculated from the onset of the fifth stimulus. Incorrect trials and RTs deviating over 2.5 SD from the mean for that particular participant were excluded from the behavioral analysis.

In the analysis of the MEG data, averaged signals were first visually inspected to identify dipolar field distributions that appeared consistently across experimental conditions and across participants. Such distributions were identified in two time windows: the M100 window (150-170 ms) and the MMF window (225-275 ms). As shown in Figure 1, the MMF, being evoked by speech stimuli, was strongly left-lateralized (an MMF response to non-speech stimuli is more evenly distributed across both hemispheres). The amplitudes and latencies of these components were recorded by calculating the root mean square (RMS) field strength from the sensors that covered the field pattern of the MMF/M100 in the left hemisphere. The sensors used for the RMS analysis were selected for each subject by creating a grand average of all 8 conditions and choosing those sensors that showed the clearest dipolar distribution and held constant across conditions within a subject. The number of sensors used varied between 4 and 6. We then compared the RMSs of the selected sensors for the response to the two oral deviant conditions to those of the corresponding standards (e.g., the *da* response following a series of *ba* to the *da* response following a series of *da*), and likewise for the nasal conditions.

4. Results

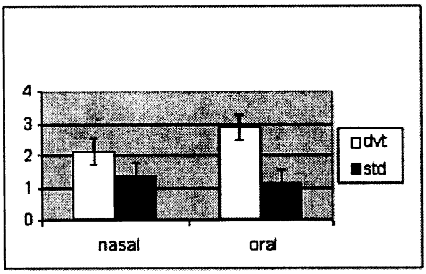
4.1. Behavioral responses

Analyses of variance (ANOVA) [(Condition[1]) X (Manner[2])] were performed for reaction times and accuracy. Deviants overall received significantly more errors

[1] Standard or deviant
[2] Oral vs. nasal

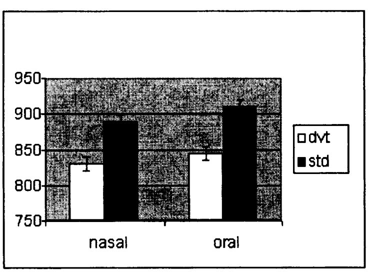
than standards (p=.0009) (Figure 3). No effect for manner was observed (p=.3538). Therefore, the results of previous behavioral studies, that nasals are more confusable than orals, were not replicated. There are several reasons why this might have been the case. First of all, the stimuli were synthesized so as to be clear and unambiguous tokens of the target syllables, to maximize the desired brain responses. No filtering was done, nor anything else to adversely affect the signal-to-noise ratio, as is typically done in behavioral confusability studies. Therefore, it is unsurprising for error rates to be unrevealing. Second, the relatively small number of trials (30 per condition) may have been a factor.

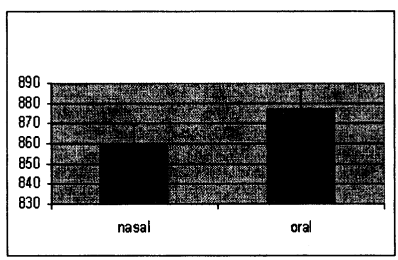
Figure 3 – Error rates. Y-axis: number of errors.



Reaction time to deviants overall was significantly faster than to standards (p≤0001). Reaction time to nasals was significantly faster than to orals (p=.0197) (Figure 4). No [(Condition) X (Manner)] effect was found. Some subjects reported that they took longer to respond to 'same' trials because they were unsure whether or not the stimulus set was complete, which could obscure reaction time patterns.

Figure 4 – Reaction times. Y-axis: time, in ms.





4.2. Electrophysiological results

4.2.1. M100 response

This response proved very difficult to locate when looking at individual subject responses. An ANOVA [(Condition) X (Manner)] was performed for amplitude and latency. No effect was found for [(Condition) X (Manner)] for either latency or amplitude, but there was a main effect for manner: overall, nasals have significantly greater amplitude than orals in the M100 time window ($p \leq .0001$).

4.2.2. MMF response

An ANOVA [(Condition) X (Manner)] was performed for amplitude and latency. No effect for latency was found. We found a main effect for amplitude for condition: deviants overall have significantly greater amplitude in the MMF time window than standards ($p \leq .0001$), as expected. An effect for manner was also found: orals overall have a significantly greater amplitude in the MMF window than nasals ($p \leq .0001$) (Figure 5). Finally, the ANOVA yielded an effect for [(Condition) X (Manner)]: as illustrated in Figure 6, the MMF/baseline gap was significantly greater for oral consonant pairs than for nasals ($p = .0399$).

Figure 5 – MMF: The following charts show a single subject's averaged response to oral deviants and standards in the top chart, and to nasal deviants and standards in the bottom chart. The MMF window (225 ms – 275 ms) shows an MMF for both orals and nasals (Y-axis: magnetic field strength, in arbitrary units proportional to femtoTesla; X-axis: time, in ms). The bar graph on the right shows the differences in peak amplitude between standards and deviants for orals and nasals.

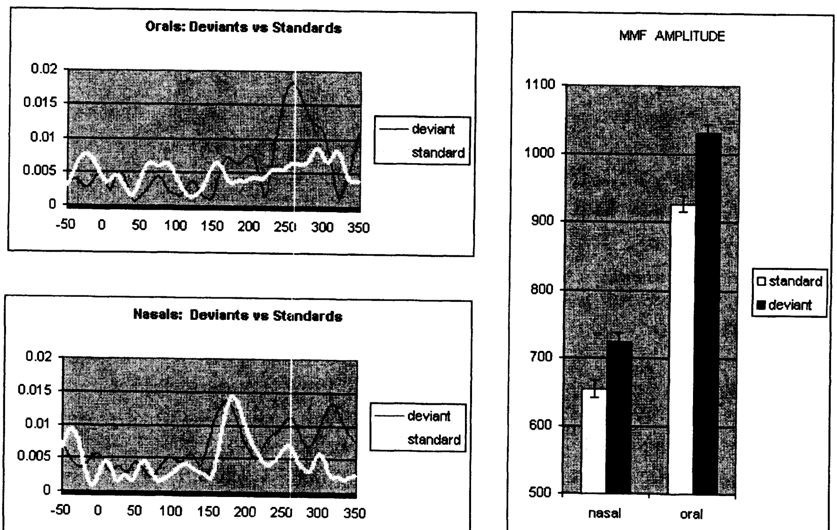
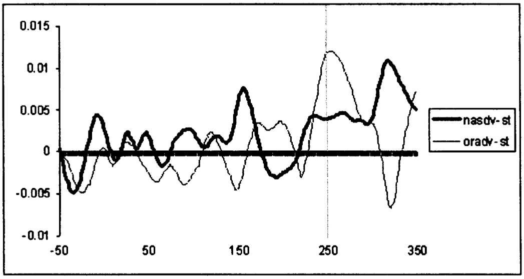
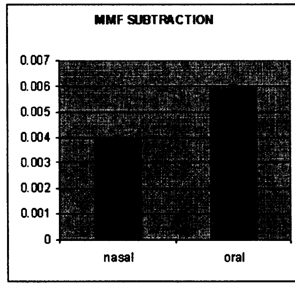


Figure 6 – MMF comparison: The following chart shows a single subject's MMF waveline (subtraction of the deviant minus standard) for nasals (thick line) and for orals (thin line). The oral waveline is higher in the MMF time window (Y-axis: magnetic field strength, in arbitrary units proportional to femtoTesla; X-axis: time, in ms). The bar graph illustrates the same subject's MMF subtraction at peak amplitude in the MMF window for nasals and orals.





5. Discussion

The predictions made by the different models of similarity discussed earlier were as follows:

- According to a perceptual similarity framework, the MMF-baseline gap (i.e., deviant minus standard) should be larger for oral consonant pairs than for nasals.
- According to a natural-class-based one such as the one presented in Frisch (1996), the gaps should be equivalent for oral and nasal pairs, or the nasal gap should be larger.
- If abstract phonological features were the only relevant factor in perceptibility at this stage, the gaps should be equivalent.

We found that the oddball paradigm evoked a stronger MMF response for orals than nasals, and thus that orals are perceived as more different from one another than nasals at this stage, the earliest available. It appears then that phonological category is not the only relevant factor in perception at this stage: acoustic similarity also plays a role. Finally, it is a perceptual similarity metric that appears to be operating at this time period, rather than a natural-class one, or feature-counting one.

It remains unclear why the M100 response was so weak for most of the subjects. However, previous studies have also reported difficulties in localizing it (Bruening et al. 2001).

6. Conclusion and future research

Our results suggest that the MMF can be used to test proposals about degree of perceptual distance in acoustic-based similarity frameworks. Our experiment showed that, once we control for phonological category, perceptual similarity matters. In future work, we would like to explore the possibility that a Frisch-type of similarity also plays a role, which is here being disguised by the stronger role of perceptual similarity, just as perceptual similarity is generally disguised by phonological category. Our next experiment will isolate Frisch-style similarity as a variable, by controlling acoustics and testing the same phonological contrast with speakers of languages that have the contrast, but whose inventories differ in other ways.

References

Bruening B., E. Flagg & V. Lin 2001 "An MEG Study of Tone Processing Asymmetries in English versus Mandarin Speakers" *Proceedings of the Northeast Linguistic Society Conference*, 31

Frisch S. 1996, "Similarity and frequency in phonology", unpublished PhD. Dissertation, Northwestern University

Frisch S., M. Broe and J. Pierrehumbert 1997 "Similarity and Phonotactics in Arabic" *Rutgers Optimality Archive 223*

Hume E. 1998 "The role of perceptibility in consonant/consonant metathesis." *Proceedings of the West Coast Conference of Formal Linguistics,* 17, 293-307

Hura S., B. Lindblom and R. Diehl 1992 "On the role of perception in shaping phonological assimilation rules," *Language and Speech* 35 (2), 59-72

Jun, J. 1995 "Place assimilation as the result of conflicting perceptual and articulatory constraints" *Proceedings of the West Coast Conference of Formal Linguistics,* 14, 221-237

Miller G. A. & P. E. Nicely 1955 "An analysis of perceptual confusions among some English consonants." *Journal of the Acoustical Society of America, 27,* 338-352

Mohanan, K. P. 1993 "Fields of attraction in phonology" in John Goldsmith (ed.) *The Last Phonological Rule*, University of Chicago Press

Näätänen R., A. Lehtoskoskl, M. Lennes, M. Cheour, M. Huotilainen, A. Ilvonen, M. Vainio, P. Alku, R. Ilmoniemi, A. Luuk, J. Allik, J. Sinkkonen & K. Alho 1997 "Language-specific phoneme representations revealed by electric and magnetic brain responses." *Nature,* 385, 432-434

Phillips C., T. Pellathy, A. Marantz, E. Yellin, K. Wexler, D. Poeppel, M. McGinnis & T. Roberts 2000 "Auditory cortex accesses phonological categories: an MEG mismatch study" *Journal of Cognitive Neuroscience,* 12

Sharma A. & M. F. Dorman 1999 "Cortical auditory evoked potential correlates of categorical perception of voice-onset time." *Journal of the Acoustical Society of America,* 106, 1078-1083

Steriade D. 1999 "P-map effects in assimilation" in Elizabeth Hume and Keith Johnson (eds.) *The role of speech perception phenomena in phonology,* Academic Press

Wang M. D. & R. C. Bilger 1973 “Consonant confusions in noise: a study of perceptual features.” *Journal of the Acoustical Society of America,* 54, 1248-1266

Wright, J. T. 1986 “The behavior of nasalized vowels in the perceptual vowel space” in J. J. Jaeger and J. J. Ohala (eds.) *Experimental Phonology*, Orlando: Academic Press, 45-67

Valentine Hacquard
Mary Ann Walter
Department of Linguistics and Philosophy
77 Massachusetts Ave, Bldg E39-245
Massachusetts Institute of Technology
Cambridge, MA 02139

hacquard@mit.edu
walterma@mit.edu

Positional markedness as a by-product of the learning situation*

ALEXEI KOCHETOV
Simon Fraser University and Haskins Laboratories

0. Introduction

This paper presents a simulation where a "learner" attempts to acquire a positionally unrestricted phonological contrast (plain vs. palatalized labial stops in a variety of contexts) without any a priori knowledge of the relevant positional hierarchies. The results show that an acquisition model that includes human-like perceptual limitations can give rise to cross-linguistically attested phonotactic grammars without positing this a priori knowledge.

1. Positional scales and learnability

Scales encoding relative markedness of environments, referred to as "positional markedness scales," play an important role in Optimality Theory (Prince & Smolensky 1993). For instance, the contrast between voiceless and voiced obstruents is known to be commonly neutralized in syllable coda position (Lombardi 1995), or, under a different analysis, in the absence of a following sonorant segment (which contains acoustic cues to the contrast: Steriade 1997). These facts can be seen as reflecting universal positional markedness scales that refer to syllable positions or phonetically-cued environments. The scales can be represented grammatically as hierarchies of positional markedness constraints (or faithfulness constraints: Beckman 1997). It is important to note that these constraint hierarchies are assumed to be *harmonic*, or universally fixed, and thus innate by definition (Prince & Smolensky 1993).

A traditional OT learner constructs a phonological grammar by ranking markedness and faithfulness constraints based on positive evidence from the lexicon s/he has acquired (Tesar & Smolensky 1998, Tesar 2000). UG supplies the learner not only with the relevant universal constraints, but also with positional (and context-free) markedness hierarchies. The question of whether the lexical items the learner is exposed to are always perfectly produced and perfectly

* Many thanks to Keren Rice, Louis Goldstein, Elan Dresher, Maria Babyonyshev, Edward Flemming, Elizabeth Hume-O'Haire, Keith Johnson, Dasha Kavitskaya, Hank Rogers, Joseph Schallert, Ron Smyth, and Donca Steriade, among others. The research was supported by fellowships from SSHRC, Canada. All errors are my own.

recoverable is hardly addressed in the OT literature on learnability (but has been raised elsewhere: Ohala 1981 et seq.).

In this paper I take a different approach to positional markedness scales, arguing that it is unnecessary to posit a priori knowledge of the scales (with or without direct reference to phonetic factors) because the attested patterns of neutralization emerge naturally from speech production and perception during language acquisition without this knowledge. The case study explored here is a positional markedness scale that reflects cross-linguistic facts of neutralization of the contrast between plain and palatalized labials.

2. A scale of distribution of the plain/palatalized contrast

A survey of over 20 languages and dialects with contrastive palatalization (Kochetov 2002) shows the following patterns: the contrast between plain (or velarized) and palatalized labials, such as /p/ vs. /p^j/, is most often restricted to the prevocalic position and is neutralized elsewhere (Type 1: Table 1); only a small number of languages allow the contrast both prevocalically and word-finally (Type 2); only one language extends the /p/ vs. /p^j/ distinction to the context before plain consonants (Type 3). A full contrast (Type 0), that is a pattern in which the contrast is expressed in all environments, appears to be unattested.[1] The distribution is also dependent on the quality of the following vowels and the place of the following consonants, which are not considered here. Note that the most common output of neutralization is a plain consonant.

Table 1. A summary of distribution of the plain/palatalized contrast in labials

Patterns	_V pa vs. p^ja	_# ap vs. ap^j	_C apta vs. ap^jta	_C^j apt^ja vs. ap^jt^ja	Comments
Type 1	yes	no	no	no	common
Type 2	yes	yes	no	no	rare
Type 3	yes	yes	yes	no	very rare
Type 0	yes	yes	yes	yes	unattested

The observed asymmetries can be represented by a scale of environments (1a) arranged from the least marked (before a vowel) to the most marked (before a palatalized consonant). Assuming that this scale is universal, that is, that it holds in all languages that maintain the plain-palatalized contrast in labials, a standard approach would be to posit a fixed hierarchy of markedness constraints that prohibit the plain/palatalized contrast (*[αpal]) in specific contexts, as in (1b) (cf., the constraints on *[αvoice] in Steriade 1997). The differences between the grammars of different language will then result from different rankings of the

[1] Some examples of the languages surveyed include Standard Bulgarian, Lithuanian, Nenets (Type 1); Russian, Irish (Type 2); and the Nova Nadezhda dialect of Bulgarian (Type 3) (see Kochetov 2002 for additional detail as well as references).

faithfulness constraint Ident[αpal] against the hierarchy of positional markedness constraints.

(1) a. __V > __# > __C > __C^j
b. *[αpal]/_C^j » *[αpal]/_C » *[αpal]/_# » *[αpal]/_V

The question, however, is whether the scale in (1) and the corresponding constraint rankings can arise without being pre-specified in UG. This possibility is explored by means of the learning simulation described below.

3. The learning situation

The approach explored here builds in part on Hayes' (2001) model of a "pure phonotactic learner." This learner, provided with a set of well-formed words in a given language and universal OT constraints, constructs a grammar (a ranking of constraints) that can evaluate well-formed and ill-formed phonotactic sequences (without reference, at this stage, to their morphological composition and alternations). At the same time, the current analysis represents a departure from other learnability work within OT. First, it assumes a more realistic speaker/listener who has inherent limitations on speech production and perception. Second, our learner is *not* provided with implicit knowledge of phonological markedness, in particular, the positional markedness scale. In other words, s/he is not aware of whether the contrast is more marked (or harder to produce/perceive) in some environments than in others (cf., Ohala 1981, Hale & Reiss 2000, Hume & Johnson 2001; cf. Mielke, this volume).[2] The current approach also relies crucially on the concept of *self-organization*, or spontaneous emergence of order, that is characteristic of many natural and artificial dynamical systems (see, for example, Kauffman 1995, Langton 1995; cf. Wedel, this volume). Specifically, in our case, the self-organization approach holds that positional markedness scales – high-level phonological structure – is a by-product of "blind" low-level interactions between speakers and listeners during the learning process.

This hypothesis with respect to the scale in (1) is investigated using a computer simulation of the interactions of agents, or simple autonomous entities. Agent-based programming provides explicit ways of testing relative importance of multiple factors, and it has been used recently to explore a variety of emergent phonological phenomena (e.g., Browman & Goldstein 1999, de Boer 2000, Liberman to appear). The current, rather simplistic, *Matlab* simulation involves two agents, an adult agent and a learning agent; both consist of components modeling speech production, perception, the lexicon, and grammar. The focus here is on the learner's ability to perceive lexical items produced by the adult, to build a lexicon, and to construct a grammar. My assumptions about the target grammar and the adult-learner interactions are presented in the following sections.

[2] The learner, however, may infer some knowledge of relative markedness (e.g., perceptibility: Steriade 2001; or articulatory complexity: Hayes 1999) from sensory/motor experience.

3.1. The hypothetical language

Let's assume a hypothetical language, Language X, that has the consonant inventory {p, p^j, t, t^j} and a vowel {a}. The lexical items of the language are presented in (2a); they consist of mono-morphemic words with a contrast between plain and palatalized labials in four environments: word-initially, word-finally, and before plain and palatalized consonants. Assume also that the knowledge of well-formed sequences of segments in this language, the phonotactic grammar, is encoded by ranking the Optimality Theory-type markedness and faithfulness constraints presented in Section 2. The grammar, that is, the ranking of these constraints characteristic of Language X, is shown in (2b). This grammar, unattested cross-linguistically (Type 0 in Table 1) will be the target grammar for our hypothetical learner.

(2) a. Lexicon: {pa, p^ja, ap, ap^j, apta, ap^jta, apt^ja, a$p^j t^j$a}
b. Grammar: Ident[αpal] » *[αpal]/_C^j, *[αpal]/_C, *[αpal]/_#, [αpal]/_V

3.2. The adult agent

The adult agent has full knowledge of the lexicon and the grammar. His/her goal in this simulation is to "produce" lexical items. This production is not directly simulated; instead, it is assumed to correspond to the actual production of the relevant consonants *at word boundaries* in nonsense utterances (Table 2). These data were obtained using an articulatory magnetometer from four speakers of Russian. It should be noted that the contrast between /p/ and /p^j/ is present in all contexts (_V, _#, _C, _C^j), although there are some gradient differences in the magnitude of tongue body raising/backing gestures and their timing relative to the lips (Kochetov, to appear). What is important here is that the adult agent's production is imperfect, having the kinds of limitations typical of a human speaker.

Table 2. Adult agent's "production"

Environments	Lexical items, Language X	Russian nonwords
(#)_V	pa, p^ja	ta [p]apy, ta [p^j]apy
(V)_#	ap, ap^j	ta[p] apy, ta[p^j] apy
(V)_C	apta, ap^jta	ta[p] tapy, ta[p^j] tapy
(V)_C^j	apt^ja, a$p^j t^j$a	ta[p] t^japy, ta[p^j] t^japy

3.3. The learning agent

Unlike the adult agent, the learning agent begins with an empty lexicon (3a), and the grammar (3b), where the markedness constraints are ranked above the faithfulness constraint (cf., Hayes 2001). The goals of the learning agent are thus (i) to recover presented items from an acoustic signal, (ii) to posit representations of these items based on a given number of tokens per item, and (iii) to rank constraints in the grammar based on generalizations over stored lexical items.

(3) a. Lexicon: {Ø}
b. Grammar: *[αpal]/_C^j, *[αpal]/_C, *[αpal]/_#, *[αpal]/_V » Ident[αpal]

Perception

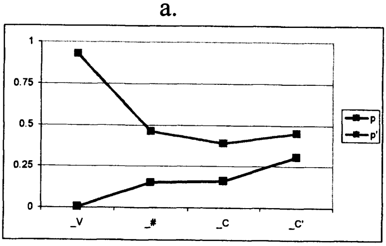
The agent's ability to recover items from the signal is limited; it is based on the probability with which humans tend to perceive the contrast /p/ vs. /p^j/ (Kochetov 2002).

The perceptibility scale used here (Figure 1a) reflects the likelihood of each segment being perceived as "palatalized" [1] or "plain" [0] in a given environment *at word boundaries* (where no phonotactic restrictions apply in Russian). It represents the average perception of the stimuli given in Table 2 under a number of experimental conditions: by native listeners without noise (N = 20), by native listeners with noise (N = 20), and by non-native listeners without noise (Japanese, N = 10).[3] A scale of the perceptual similarity of /p/ and /p^j/ in different phonological contexts (Figure 1b) was computed from confusion matrices (based on correct responses and false alarms; see Johnson 2003). Note that the perceptual scales are assumed here to reflect universal performance biases in auditory perception that are external to the phonological grammar proper (cf., Hume & Johnson 2001).

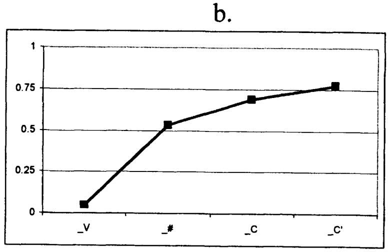
As we can see in Figure 1, both the "plain" or "palatalized" responses and the degree of similarity between the segments are substantially affected by environments: there is an almost 100% correct identification rate of /p/ and /p^j/ in the prevocalic context, but a high rate of confusion word-finally and before consonants (especially for /p^j/).[4] The positional effect is striking, especially given the fact that all these word-boundary contexts are phonotactically possible both for native and non-native listeners (the Japanese listeners interpreted coda [p] and [p^j] as syllables /pu/ and /pi/).

Figure 1. Mean perception of /p/ and /p^j/ as "plain" [0] or "palatalized" [1] in four contexts at word boundaries (a); mean perceptual similarity between /p/ and /p^j/ in four contexts (b); "0" is "different," "1" is "same."

a.



b.



[3] Non-native listeners were used as an additional control for language-specific biases.
[4] Articulatory and acoustic factors that may have contributed to these results are discussed in Kochetov (2002).

To model the likelihood of the consonants being perceived as "plain" or "palatalized," the simulation employs a random number generator. For instance, if a random number from 0 to 1 is less than 0.93 (see Figure1a) the item {p^ja} is "perceived" by the listener/learner as {p^ja} ("palatalized"). If the number is greater than 0.93, the item is "perceived" as {pa} ("plain"). The perceived tokens – whether correctly identified or confused with other forms – are stored in memory. In each of the tokens the consonant is specified as "palatalized" by assigning it the value [1], or "plain," by assigning it the value [0].

Lexicon and grammar

Based on the stored tokens, the learner posits a single form for a given lexical item, that is, its representation. The representation of the form is determined based on the following labeling algorithm: If the mean for all the stored tokens for a given item has a score of less than 0.45 (i.e., most tokens are "plain"), then the item is judged to have a non-palatalized consonant, [0]. If the score is higher than 0.55 (i.e., most tokens are "palatalized"), the item is judged to have a palatalized consonant, [1]. If the item falls within the range 0.45-0.55, it is randomly assigned one of the two labels. For example, if 56 out of 100 tokens of {ap^j} are perceived as "plain" (an average of 0.56) and 44 tokens are perceived as "palatalized" (an average of 0.44), the form is labeled, or "saved" in the lexicon, as {ap^j}. A lexicon is represented as a matrix of ones and zeros: for example, [0 0 0 1; 1 0 0 1] represents a lexicon containing the items {pa, ap, apta, ap^jt^ja; p^ja, ap, apta, ap^jt^ja}. These initial lexical representations are updated based on the learner's perception of his/her own production of pairs of items (e.g., {apt^ja} and {ap^jt^ja}). If the pairs of items are perceived as being similar, their representations are revised (e.g., to either {apt^ja} or {ap^jt^ja}). The two pairs are considered similar if a generated random number is below the perceived similarity score for this pair (Figure 1b).

Based on the lexicon acquired, the learner constructs the grammar by ranking the constraints. Recall that the initial state of the learning agent is the ranking in (3b): the faithfulness constraint is ranked at the bottom of the hierarchy and the positional markedness constraints are not pre-ranked with respect to each other. This is important, because the constructed grammar becomes crucially dependent on the items the learner stores in the lexicon. In the case of perfect perception the agent's lexicon and grammar would be an exact reflection of the target lexicon and grammar. If the learner's perception is completely random, we may expect a wide range of possible lexicons and grammars (16 logically possible permutations of the constraints in (3b)).

In sum, the processing of the presented items by the learner thus consists of recovering these items and making several types of simple generalizations: generalizations over tokens of an item (lexical representations), over pairs of lexical items (lexical contrast), and over the lexicon in general (the grammar). Although the learning situation described here is a much simplified model of a real-life acquisition path, the simplifications will allow us to focus on the phenomenon under investigation, as well as to keep the simulation computationally tractable.

Having presented the mechanism of the simulations, I turn to the following question: Is our "imperfect" learner capable of developing the target lexicon and constructing the target grammar, and if not, what are the resulting lexicons and grammars?

4. Results

In this section I describe the results of the simulation representing a sample run based on 100 tokens per item and 100 iterations. This means that our learner is presented each time with 100 tokens of 8 lexical items from Lexicon 0. The learner recovers these tokens and posits 8 representations. The representations are updated after 100 productions of pairs of items and these updated representations then constitute the learner's lexicon. Based on the items in the lexicon, the learner ranks 5 constraints, constructing the grammar. This process is repeated 100 times (with the same input, Lexicon 0), that is, one run of the simulation produces 100 grammars. After reviewing the results, presented separately for the lexicons (Section 4.1) and the grammars (Section 4.2), I address the issue of sensitivity of the results to parameter settings (Section 4.3) and relevance of other factors (Section 4.4).

4.1. Building the lexicon

Table 3 summarizes the results of a sample run in terms of major lexicon types; any deviations from the target lexicon are shaded. Lexicon 0 is the target lexicon that that contains 8 phonologically distinct lexical items. Note that this lexicon is not once replicated by the learner (0 occurrences). Lexicon 1 shows a high degree of homophony, with {ap, apta, apt^{j}a or apt^{j}a} corresponding to the original pairs {ap vs. ap^j, apta vs. ap^{j}ta, apt^{j}a vs. ap^{j}t^{j}a}. This lexicon is by far the most common outcome: it was generated 74 out of 100 times.[5] Lexicon 2, which differs from the previous lexicon in preserving the original distinction {ap vs. ap^j} was produced by the learner only 12 times. Lexicon 3 is listed here because it is minimally different from the target Lexicon (by the lack of the contrast {apt^{j}a vs. ap^{j}t^{j}a}); this lexicon occurs in the output only once. It is thus as likely as (or even less likely than) some other generated lexicons, grouped under the title "Other lexicons". In the current sample these include three lexicons that collapse the distinctions between {ap vs. ap^j} or {ap^{j}ta vs. ap^{j}t^{j}a} in favor of the palatalized forms {ap^j ap^{j}ta}.[6] Note that all of the resulting lexicons keep the distinctions between the items {pa} and {p^{j}a}.

[5] The results for the last environment (_C^j) are combined here; there were 57 occurrences of {apt^{j}a} and 17 occurrences of {ap^{j}t^{j}a}.

[6] This result is due to the learner's random selection of one of the relatively similar forms (see Section 3.3).

It should be mentioned that additional runs under the same conditions showed very similar results: the absence of Lexicon 0, and relatively high frequency of Lexicons 1 and 2. Overall, the degree of variation among the lexicons generated is quite limited, with certain types of lexicons clearly preferred over others.

The results of the simulation are not surprising given the learner's human-like perceptual limitations and biases: The learning agent tends to perceive the item {ap^j} as [ap] almost as often as [ap^j]; the same, however, does not hold for the item {ap} (see Figure 1). The pairs {apta} vs. {ap^{j}ta} and {apt^{j}a} vs. {ap^{j}t^{j}a} have an even higher error rate and perceived similarity. Generalizations over tokens and pairs of items help the learner select one form or the other, but they do not eliminate the strong biases towards certain forms (e.g., {ap} or {apta}).

Table 3. Resulting lexicons: A sample run based on 100 tokens (100 iterations)

<table>
<tr><td colspan="2">Lexicon 0</td><td colspan="2">Lexicon 1</td><td colspan="2">Lexicon 2</td><td colspan="2">Lexicon 3</td><td colspan="2">Other lexicons</td></tr>
<tr><td colspan="2">0 occurrences</td><td colspan="2">74 occurrences</td><td colspan="2">12 occurrences</td><td colspan="2">1 occurrence</td><td colspan="2">13 occurrences</td></tr>
<tr><td>pa</td><td>p^{j}a</td><td>pa</td><td>p^{j}a</td><td>pa</td><td>p^{j}a</td><td>pa</td><td>p^{j}a</td><td>pa</td><td>p^{j}a</td></tr>
<tr><td>ap</td><td>ap^j</td><td colspan="2">ap</td><td>ap</td><td>ap^j</td><td>ap</td><td>ap^j</td><td colspan="2">ap or ap^j</td></tr>
<tr><td>apta</td><td>ap^{j}ta</td><td colspan="2">apta</td><td colspan="2">apta</td><td>apta</td><td>ap^{j}ta</td><td colspan="2">apta or ap^{j}ta</td></tr>
<tr><td>apt^{j}a</td><td>ap^{j}t^{j}a</td><td colspan="2">apt^{j}a or ap^{j}t^{j}a</td><td colspan="2">apt^{j}a or ap^{j}t^{j}a</td><td colspan="2">apt^{j}a</td><td colspan="2">apt^{j}a or ap^{j}t^{j}a</td></tr>
</table>

4.2. Constructing the grammar

Having posited the lexical forms, the learner constructs a subset of the phonotactic grammar, a set of generalizations about how segments can be combined in the language. This is done by demoting positional markedness constraints based on evidence from the lexicon (e.g., Hayes 2001).

Unsurprisingly, our results reveal zero occurrences of the target grammar, as shown in Table 4. Recall that, in this least restrictive grammar, the faithfulness constraint Ident[□pal] is ranked above all the positional markedness constraints. The set of grammars the learner does arrive at is very limited. The most common outcome by far is Grammar 1, which neutralizes the contrast in all contexts except the prevocalic one. Note that, for simplicity, the constraints used do not specify whether the outcome of neutralization is a plain or a palatalized segment (but see Kochetov 2002). Thus, Grammar 1 can be constructed based on any lexicon that preserves the original distinction among {pa} vs. {p^{j}a}, but not between other items (Lexicon 1 and the lexicons listed in the last column). Grammar 2, a consequence of positing Lexicon 2, is much less common. Finally, given the unlikely Lexicon 3, the corresponding grammar (Grammar 3) is also a rare outcome. Additional runs under the same conditions replicate the strong bias towards Grammars 1 and 2.

Note that we do not find in the results a number of other logically possible rankings (in fact, 12 out of 16 logically possible rankings). These include, for instance, grammars that disallow the contrast word-initially while maintaining it word-finally or preconsonantally. Given the probabilistic nature of the learner's perception, there is no categorical prohibition against these rankings (as well as the target grammar) — these grammars are theoretically possible, yet very unlikely due to the learner's perceptual abilities. It should be noted, however, that other factors (discussed below) can override the substantive biases.

Table 4. Resulting grammars: A sample run based on 100 tokens (100 iterations)

Grammar 0	Grammar 1	Grammar 2	Grammar 3	Other grammars
0 occurrences	87 occurrences	12 occurrences	1 occurrence	0 occurrences
Ident[αpal] » *[αpal]/_Cʲ, *[αpal]/_C, *[αpal]/_#, *[αpal]/_V	*[αpal]/_Cʲ *[αpal]/_C, *[αpal]/_#} » **Ident[αpal]** » *[αpal]/_V	*[αpal]/_Cʲ *[αpal]/_C} » **Ident[αpal]** » *[αpal]/_#, *[αpal]/_V	*[αpal]/_Cʲ » **Ident[αpal]** » *[αpal]/_C, *[αpal]/_#, *[αpal]/_V	...

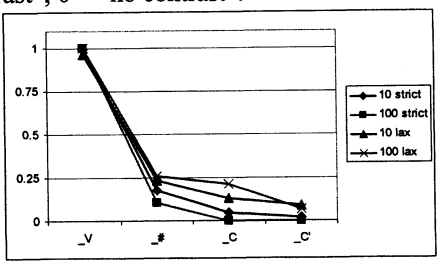
The most significant outcome of our simulation is that the results reflect in general the cross-linguistically attested patterns of neutralization of the plain-palatalized contrast (see Table 1). Thus the most commonly generated grammar corresponds to the most commonly attested pattern of neutralization of the plain-palatalized contrast in labials (Type 1). Grammar 2 is a less common outcome; it is also less common cross-linguistically (Type 2). Finally, the very unlikely outcome, Grammar 3, is attested in only a single language. The simulation thus makes correct predictions about the range of cross-linguistic variation with respect to the neutralization of the contrast, as well as to the relative frequency of each particular pattern. Crucially, to the analysis that the positional markedness scale (_V > _# > _C > _Cʲ) has been achieved without or independently of the learner having prior knowledge of it.

4.3. Parameter settings

The question we need to address next is to what extent the results discussed in the previous sections are influenced by the parameters specified in the current simulation. Recall that our learner was presented with 100 tokens of each lexical item. Lexical representations were posited based on the following, quite deterministic, labeling algorithm: *if* the mean score x ∈ 0-0.45 *then* "plain", *if* x ∈ 0.55-1 *then* "palatalized", *if* x ∈ 0.45-0.55 *then* either of the two (see Section 3.3). Subsequent manipulations of these parameters have shown them to affect the results to a certain degree, but within important limits.

Figure 2 plots the mean contrast between /p/ and /pʲ/ in each environment (based on 100 iterations) resulting from four different settings: either 10 or 100 tokens categorized by the algorithm described above ("10 strict" and "100 strict"); either 10 or 100 tokens categorized by a less deterministic labeling algorithm ("10 lax" and "100 lax"): *if* the mean score x ∈ 0-0.25 *then* "plain", *if* x ∈ 0.75-1 *then* "palatalized", *if* x ∈ 0.25-0.75 *then* either of the two. Note that under all four settings the simulation results in grammars all of which maintain the contrast prevocalically (with 3 exceptions under "10 lax"). The differences are primarily in the degree of variability (maintained or neutralized) in the less favorable environments (_#, _C, and C^j); with all of the settings, however, the contrast is supported in a relatively small number of grammars. Overall, fewer tokens and less deterministic labeling procedures lead to somewhat higher variability in the output.

Figure 2. Resulting contrasts between /p/ and /p^j/ in 4 environments based on 10 or 100 tokens and "strict" or "lax" labeling algorithms; the mean for 100 iterations; 1 = "contrast", 0 = "no contrast".



4.4. Limitations and additional factors

A number of questions related to the results of the simulation require further consideration. First, the current simulation did not involve the possibility of feedback from the adult agent (see, e.g. de Boer 2000). This factor would undoubtedly increase the learner's chances of attaining the target grammar. Morphological decomposability of lexical items would also contribute to the maintenance of the contrast in less favorable environments, reinforced by other kinds of lexical/grammatical generalizations, specifically, output-output correspondence relations between allomorphs (Benua 2000; see Kochetov 2002). Implementation of these factors is currently under way. Further, positing an a priori set of context-specific constraints made the simulation more manageable, but it would be more desirable for the learner to infer these from the input (cf., Hayes 1999). Further work should aim to rely on more complex interactions in agent populations over time and on a model of human (child and adult) speech production and perception that is more realistic and language-independent. Thus, one of the limitations of the current simulation is in its reliance on language-particular production of palatalized labials (Flemming, p.c.; see Kochetov 2002 for a discussion). This problem can be addressed through the use of both synthesized speech and data from a number of languages. Also, the current simulation does not predict the commonly attested contrast enhancement phenomenon (Flemming 1995). This effect can possibly result from continuous multiple-agent interactions (cf., de Boer 2000) with meaningful lexical items used in more natural communication environments. Finally, a more realistic simulation should give more attention to higher-level lexical and grammatical generalizations (see Wedel, this volume) and should use a wider range of lexical items and phonological contrasts.

5. Conclusion

In sum, the goal of the simulated learning situation was to provide us with insight into the nature and the mechanism of the emergence of positional markedness asymmetries in palatalization, and in general. An important result of the simulation is that it is unnecessary to attribute an innate status to positional markedness scales, or perhaps even to markedness scales in general. These phonological structures can arise and evolve through relatively simple and repetitive speaker-listener interactions during the course of language acquisition.

References

Beckman J. 1997. Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14: 1-46.

Benua, L. 2000. *Phonological relations between words.* New York, N.Y.: Garland Publishing.

Browman, C. P., and Goldstein, L. 1999. Competing constraints on intergestural coordination and self-organization of phonological structures. In *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5. Grenoble: 25-34.

de Boer, B. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28: 441-465.

Flemming, E. 1995. Auditory representations in phonology. PhD dissertation, UCLA.

Hale, M., and C. Reiss. 2000. Substance abuse and dysfunctionalism: Current trends in phonology. *Linguistic Inquiry* 31: 157-169.

Hayes, B. 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer, and K. Wheatly, eds., *Functionalism and Formalism in Linguistics*, Vol. I: General Papers, John Benjamins, Amsterdam: 243-285.

Hayes, B. 2001. to appear. Phonological acquisition in Optimality Theory: The early stages. In *Fixing priorities: Constraints in phonological acquisition*, eds. R. Kager, J. Pater, and W. Zonneveld. Cambridge: Cambridge University Press.

Hume, E., and K. Johnson. 2001. A model of the interplay of speech perception and phonology. In *The role of speech perception in phonology*, eds. E. Hume and K. Johnson. New York, N.Y.: Academic Press: 3-26.

Johnson, K. 2003. *Acoustic and auditory phonetics.* 2nd edition. Cambridge, MA: Blackwell.
Kauffman, S. 1995. *At home in the universe: The search for the laws of self-organization and complexity.* Oxford: Oxford University Press.
Kochetov, A. 2002. *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization.* New York, London: Routledge.
Kochetov, A. to appear. Syllable position effects and gestural organization: Articulatory evidence form Russian. In *Papers in Laboratory Phonology VIII,* Berlin, New York: Mouton de Gruyter.
Langton, C. 1995. *Artificial Life: An overview.* MIT Press, Cambridge, MA.
Liberman, M. to appear. Simple models for emergence of a shared vocabulary. In *Papers in Laboratory Phonology VIII,* Berlin, New York: Mouton de Gruyter.
Lombardi, L. 1995. Laryngeal neutralization and syllable well-formedness. *Natural Language and Linguistic Theory* 13: 39-74.
Mielke, J. this volume. The diachronic influence of perception: Experimental evidence from Turkish. *Proceedings of the 29th meeting of the Berkeley Linguistic Society* (BLS 29).
Ohala, J.J. 1981. The listener as a source of sound change. In *Papers from Chicago Linguistic Society parasession on language and behavior,* C.S. Masek, R.A. Hendrick, and M.F. Miller. Chicago, IL: CLS: 178-203.
Prince, A., and P. Smolensky 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University.
Steriade, D. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms., University of California, Los Angeles.
Steriade, D. 2001. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Ms.
Tesar, B., and P. Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29: 229-268.
Tesar, B. 2000. *Learnability in Optimality Theory.* Cambridge, MA: MIT Press.
Wedel, A. this volume. Self-organization and the origin of higher-order phonological patterns. *Proceedings of the 29th meeting of the Berkeley Linguistic Society* (BLS 29).

Alexei Kochetov
Department of Linguistics
Simon Fraser University
Burnaby, British Columbia
V5A 1S6 Canada

alexei_kochetov@sfu.ca

The Diachronic Influence of Perception: Experimental Evidence from Turkish

JEFF MIELKE[1]
The Ohio State University

1. Phonetics in Phonology: Conflicting Hypotheses

Two conflicting strategies have been proposed to explain the role of phonetic factors such as speech perception in phonology. According to the "Sound Change" view, phonetic explanation should be invoked solely in explaining diachronic changes which give rise to synchronic alternations (see, e.g., Ohala (1981, 1993), Anderson (1981), Dolbey and Hansson (1999), Buckley (2000), Hale and Reiss (2000), Hume and Johnson (2001), Hyman (2001), Janda and Joseph (2002), Myers (2002)). According to the "Direct Influence" view, phonetic explanation should also be incorporated into synchronic phonology (see, e.g., Archangeli and Pulleyblank (1994), Pater (1996), Hayes (1997), Steriade (1997, 2001, this volume), Flemming (1995), Myers (1997), Mielke (2000)). The purpose of this study is to test some of the predictions made by these two approaches.

An example of Direct Influence is the P-map hypothesis (Steriade 2001, this volume), which states that speakers' knowledge of the perceptibility of contrasts influences phonology directly: "The primary function of the P-map is to guide the speaker in search of the minimal input deformation that solves a phonotactic problem. The grammatical reflex of the P-map is the projection of and ranking among correspondence constraints" (Steriade, In press:4). An alternative explanation, argued by e.g., Ohala (1981, 1993 *inter alia*), Hume and Johnson (2001), and Hyman (2001), is that perception is relevant at the inception of a sound change, but it is not part of the grammar.

[1] Thanks to Mike Armstrong, Savas Arslan, Mary Beckman, Robin Dodsworth, Shelome Gooden, Tsan Huang, Beth Hume, Brian Joseph, Keith Johnson, Alexei Kochetov, Kristie McCrary, Grant McGuire, Scott Myers, Dave Odden, Phonies, Janet Pierrehumbert, Charles Reiss, Anton Rytting, Misun Seo, Andrea Sims, Donca Steriade, the creators of TELL, Giorgos Tserdanelis, Pauline Welby, Steve Winters, Alan Yu, audiences at OSU, at the 2002 MOT Phonology Workshop in Montréal, and at the GDR Phonologies 4èmes Journées Internationales in Grenoble, and people at poster sessions at OSU's CogFest and LabPhon 8, for valuable assistance with this paper, including many helpful comments and suggestions.

Both approaches seek to explain why many formally statable phonological processes fail to occur (The Too-Many-Solutions problem). Steriade's solution to the Too-Many-Solutions problem is that a speaker's knowledge of perceptual similarity projects a ranking of correspondence constraints, and this constraint ranking *predicts* attested patterns. An alternative solution to the Too-Many-Solutions problem is that perceptibility constrains the set of likely grammars diachronically. Perception and other phonetic factors may motivate a sound change at its inception (see Janda and Joseph 2002), but do not play an active role in the synchronic grammar. Because perception contributes to the likelihood of a sound change, its influence is seen in synchronic phonological alternations.

Some observations are consistent with the P-map hypothesis. For example, there is a crosslinguistic preference for less noticeable repair strategies (ones which involve maximally similar UR and surface forms), and speakers are aware of similarity rates, based on experimental evidence and on studies of rhyming patterns (Steriade, this volume, and references cited). Steriade concludes that the preference for less noticeable repair strategies is teleological, that speakers choose less noticeable repairs because they know they are less noticeable. Even if speakers possess similarity knowledge, this is not direct evidence that they apply it, so empirical evidence is needed.

Proving the P-map hypothesis requires more than just evidence that perception influences phonology (because the influence could be diachronic) and that speakers have knowledge of similarity (because this knowledge may not be applied to phonology). It requires testing whether perceptibility knowledge is synchronically accessible to and used by speakers. Because it assumes that perception influences phonology via speakers' adaptive strategies (following, e.g., Lindblom 1990), the P-map hypothesis predicts that perception could manifest itself in speech style variation as speakers adapt to different situations (Steriade, p.c.). The sound change hypothesis does not make this prediction, so a close examination of a perceptually-conditioned process may be informative. Turkish /h/ deletion is one such process.

2. Turkish /h/ Deletion: Divergent Predictions

/h/ is optionally deleted in Turkish casual speech, but only in certain segmental contexts (Lewis 1967, Sezer 1986). Sezer reports that /h/ is optionally deleted *before* sonorant consonants (1a), but not after them (1b). /h/ is optionally deleted *after* voiceless stops (2b) and affricates (3b), but not before them (2a & 3a). /h/ is optionally deleted before *and* after voiceless fricatives (4a & 4b), and it is optionally deleted intervocalically (5a), but not word-initially (5b). Sezer reports that /h/ does not delete word-finally (5c), but informal native speaker judgments indicate that it deletes in this environment as well, and a previous production experiment shows that word-final /h/ deletion is conditioned by the initial segment of the following word, occurring under conditions similar to those under which word-internal /h/ is deleted (Mielke 2002a).

(1)	a.	fihrist	~	fi:rist[2]	'index'
		tehlike	~	te:like	'danger'
		mehmet	~	me:met	proper name
		köhne	~	kö:ne	'old'
	b.	merhum		*merum	'the late'
		ilham		*ilam	'inspiration'
		imha		*ima	'destruction'
		tenha		*tena	'deserted'
(2)	a.	kahpe		*ka:pe	'harlot'
		sahte		*sa:te	'counterfeit'
		mahkum		*ma:kum	'inmate'
	b.	ʃüphe	~	ʃüpe	'suspicion'
		ethem	~	etem	proper name
(3)	a.	ahtʃi		*a:tʃi	'cook'
	b.	metʃhul	~	metʃul	'unknown'
(4)	a.	mahsus	~	ma:sus	'special to'
		tahsil	~	ta:sil	'education'
		ahʃab	~	a:ʃab	'made of brick'
	b.	ishal	~	isal	'diarrhea'
		safha	~	safa	'step'
		meʃhur	~	meʃur	'celebrity'
(5)	a.	tofium	~	toum	'seed'
		müfiendis	~	müendis	'engineer'
		safian	~	saan	'copper food dish'
		mufiafaza	~	muafaza	'protection'
	b.	hava		*ava	'air'
	c.	timsah		?timsa:	'crocodile'

A perception experiment (Mielke 2002b, To appear) indicates that /h/ is generally less perceptible in environments where it deletes than in environments where it does not delete. Crucially, for present purposes, perceptibility varies substantially between different deletion environments, as seen below in (6). Because /h/ deletion is optional and perceptually conditioned, the different approaches make different predictions about how the deletion rates in relatively high and relatively low perceptibility environments will be affected by speech style variation.

[2] When it is deleted from preconsonantal or final position, compensatory lengthening of the preceding vowel occurs.

(6) Ability of Turkish listeners to perceive /h/ in its deletion environments (Mielke 2002b, To appear)

Context	**Sensitivity (d')**	
V __ glide	3.462	most salient
V __ liquid	3.340	
V __ nasal	3.303	
intervocalic	2.897	
V __ voiceless fricative	2.663	
voiceless affricate __ V	2.633	
voiceless stop __ V	2.408	
voiceless fricative __ V	2.300	
word-final	1.123	least salient

The P-map hypothesis assumes that Turkish speakers are aware of differences in perceptibility, and that they delete /h/ in environments where a change is less likely to be noticed. A further prediction is that in more careful speaking styles, as compared to more reduced styles, /h/ will delete more frequently in environments where deletion is less noticeable. In a production experiment with /h/ deletion rate as a dependent variable, this should appear as a significant interaction between speaking style and segmental context.

According to the sound change hypothesis, differences in perceptibility are synchronically irrelevant, because only two arbitrary categories are relevant: the environments where deletion occurs and the environments where deletion does not occur. This predicts that speech style variation will affect deletion equally in all environments, and in careful styles, /h/ deletion will not be more skewed toward low-perceptibility environments than it is in reduced styles. Thus, there should be no significant interaction between speaking style and context.

3. Production Experiment

3.1. Methods

A production experiment was designed to test the predictions in the previous section. In order to induce speech style variation and /h/ deletion, 100 target words containing /h/ in various segmental contexts were elicited from native Turkish speakers in a word list, in a story, and in a story with a cognitive load. Subjects were nine female and one male native speakers of Turkish in Columbus, Ohio, aged 21-34. 100 target words with /h/ in the segmental contexts in (7) were selected from the TELL[3] database. The target words were elicited from subjects three times and recorded in mono with a Shure SM10A head-mounted

[3] The Turkish Electronic Living Lexicon (http://socrates.berkeley.edu:7037/TELLhome.html) was developed by Sharon Inkelas with John Lowe, Aylin Kuntay, Ronald Sprouse, Yelda Mesbah and Orhan Orgun.

microphone, a Symetrix SX202 dual mic preamp, and a Sony DTC-790 DAT recorder. The recordings were then digitized at 44100 Hz using Syntrillium's CoolEdit audio editing software.

(7) Target words

Deletion Environment		#	Non-deletion Environment		#
voiceless stop __ V	[p, t, k]	10	V __ voiceless stop	[p, t, k]	10
V __ nasal	[n, m]	10	nasal __ V	[n, m]	10
V __ liquid	[l, r]	10	liquid __ V	[l, r]	10
V __ voiceless fricative	[f, s, ʃ]	10	word-initial		10
voiceless fricative __ V	[f, s, ʃ]	10			
intervocalic		10	TOTAL # OF TOKENS		100

The instructions for each successive production of the target words were designed to produce different speaking styles. Subjects read the 100 target words in a word list with 250 filler words in order to produce careful speech. Subjects also read the same target words in a story, which was intended to induce more reduced speech. Subjects read the story a second time with a cognitive load, to induce further speech reduction by distracting subjects from the speaking task (Harnsberger & Pisoni 1999). For the cognitive load, subjects memorized an 8-digit number before the first, third, and fifth pages of the six-page story and attempted to repeat each number after reading the next two pages. To avoid lexical priming by the word list, the story-reading tasks were performed before the word list task.

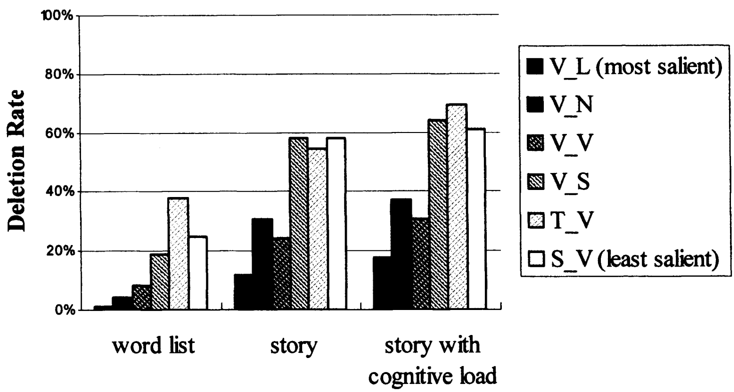
The 3000 target words were spliced out of context and separated into two groups according to whether /h/ was retained or deleted, according to the following criteria. If the word contained an audible /h/, deletion did not occur. If /h/ was inaudible or ambiguous, spectrograms were created, and if the spectrogram indicated any sign of an /h/ gesture, i.e., voicelessness or frication not attributable to another consonant, deletion did not occur. If the spectrogram indicated no sign of an /h/ gesture, deletion did occur.

3.2. Results and Analysis

For the 60 target words with /h/ in environments where Sezer (1986) reports deletion (8), /h/ deletion rates were 15.8% (95/600) for the word list task, 39.5% (237/600) for the story task, and 46.5% (279/600) for the story with a cognitive load. For the 40 target words with /h/ in environments where deletion is not reported, /h/ deletion rates were 3.5% (14/400) for the word list task, 19.3% (77/400) for the story task, and 26.0% (104/400) for the story with a cognitive load. Pairwise *t*-tests indicated no significant differences between deletion rates in any of the environments where deletion is not reported.

A univariate analysis of variance (ANOVA) was performed on the arcsine transforms of the deletion rates in the six deletion environments, and found main effects for **task** [$F(2,162) = 49.043$; $p < 0.001$] and **environment** [$F(5,162) = 24.698$; $p < 0.001$], but no significant interaction between **task** and **environment** (which is predicted by the P-map hypothesis) [$F(10,162) = 0.532$]. A second ANOVA was performed for just the 'story' and 'story with cognitive load' tasks, and found a main effect for **environment** [$F(5,108) = 16.722$; $p < 0.001$], a near-effect for **task** [$F(1,108) = 3.562$; $p = 0.062$], and again, no significant interaction between **task** and **environment** [$F(5,108) = 0.094$].

(8) Deletion rates (V = vowel, L = liquid, N = nasal, S = voiceless fricative, T = voiceless stop, # = word boundary)



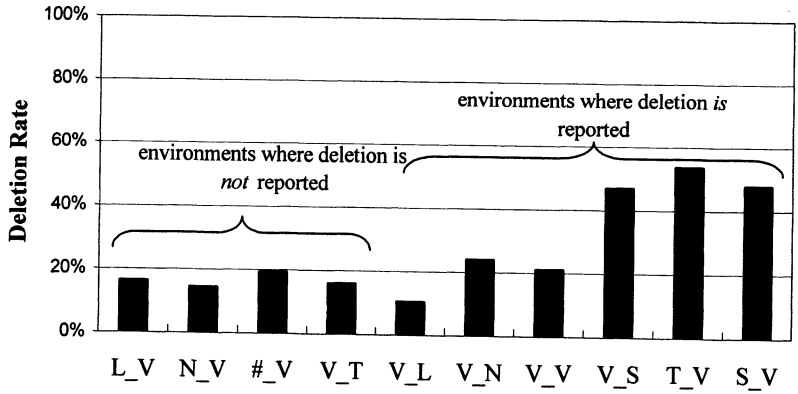
Deletion rates are lower in the three most salient environments (V_L, V_N, and V_V), but this is not limited to any particular speaking style, as predicted by the P-map hypothesis. The reason for this becomes clear upon examining the results for the four environments where deletion is not claimed to occur. Some deletion occurs in all environments, including the four environments where deletion is not reported by Sezer (1986), and the deletion rates in three of the environments where deletion *is* reported are not significantly higher. Significantly higher rates of deletion exist only for three environments (V_S, T_V, and S_V).

Comparing the six deletion environments with the environments where deletion is not reported to occur reveals that deletion rates are particularly high only for three of the reported deletion environments (9).[4] Pairwise *t*-tests show

[4] Another sensible place to look for explanation is the frequency of words with /h/ in these contexts. A search of a newspaper corpus of about 350,000 words gathered by Kemal Oflazer in the late 1990s shows no correlation between deletion environments and frequency of occurrence of words which contain /h/ in these environments. Of the environments in (9), by far the most frequent in both token and type frequency are word-initial /h/ and intervocalic /h/. Among the consonant environments, two of the environments with high deletion rates are very low in

that the deletion rates before and after voiceless fricatives and after voiceless stops are significantly higher than in each of the other environments [$p < .002$]. The clustering of environments into two categories suggests the existence of more than one deletion process in Turkish, something which is addressed below.

(9) Deletion rates in all environments (averaged across three tasks)



4. Discussion

The results support the sound change hypothesis. The lack of a significant interaction between **task** and **environment** means that the distribution of deletions does not change according to speaking style. The overall rate of deletion does indeed increase as speech style becomes more casual, but the slope does not change significantly. This means that perceptibility does not appear to be relevant for determining where deletions will occur in these speaking styles. The results of the experiment provide no evidence that perception is part of the grammar.

The results do provide further support for the role of phonetic explanation in phonology, because the deletion pattern found in this study is very much in line with the predictions of a perceptual account. In half of the environments where deletion is reported by Sezer (1986), the deletion rates are virtually the same as in environments where deletion is not reported. Deletion rates are significantly higher only before and after voiceless fricatives and after voiceless stops. The deletion rates in these environments form a homogenous subset apart from the other seven, and they also share two properties which make them likely candidates for conditioning /h/ deletion. First, /h/ is less salient to Turkish listeners in these environments than in any other environment included in the study, due ostensibly to the presence of masking frication and aspiration noise

frequency, but /h/ after voiceless stops is more common than all but two of the other consonant environments.

which is adjacent to /h/ in these environments but either nonexistent or not adjacent to /h/ in the other seven (Mielke 2002b, To appear). Second, /h/ is adjacent to a segment with laryngeal opening and closing gestures in these three environments (and also before voiceless stops). These environments are more prone to articulatory simplification via aggregation of laryngeal gestures, as seen in the results of Munhall and Löfqvist's (1992) "Kiss Ted" experiment. This is further evidence for the importance of perception and production factors as *external* forces which shape phonological systems.

Based on deletion rates found in this study, one possible conclusion is that the deletion process described here is not the same as the one described by Sezer. This study finds a nominal rate of deletion in all environments, and a smaller set of environments where a distinct deletion process apparently applies. The differences between the two sets of findings may be attributable to regional variation, generational differences, change in the nature of the /h/ deletion process over the course of 25 years, or simply the fact that Sezer's study is based on native speaker intuitions and this one is based on experimental data.[5]

The results of this study suggest that the Too-Many-Solutions problem has already been solved in language change. If a process is already active in a language, the Too-Many-Solutions problem is never even posed to the language learner or language innovator. This is consistent with Myers' (2002) proposed solution to the problem of gaps in the factorial typology of Optimality Theory. Myers (2002:28) argues that gaps in factorial typology are "pervasive and natural", because the alternations which do not occur are those that cannot arise through simple sound changes from phonetic patterns.

This explanation is similar to Steriade's in that both are explanatory on functional grounds. The difference lies in whether or not the functional grounding which transparently interacts with language change should be included in the synchronic phonology as well. Steriade has demonstrated that inclusion of perceptual knowledge in the phonological grammar is possible, because speakers have perceptual similarity judgments, and because languages frequently employ relatively non-salient repair strategies. However, it has not been shown that speakers employ similarity judgments in their phonological processing. This study of the interaction of phonology with speech style variation finds no evidence for such a synchronic relationship between functional factors and phonology. That said, there are a number of ways in which this argument against the P-map hypothesis can be shown to be wrong.

First, one could argue that deletion is not the best example of a listener-oriented phonological process, and that the lack of evidence for listener-oriented adaptive strategies in Turkish /h/ deletion does not rule out the possibility of finding such strategies in fortition processes or in hyperarticulated speech.

[5] Repetition of the above ANOVAs with the seven low-deletion environments and the three high-deletion environments analyzed separately also found no significant interaction, although in this case the set of high-deletion environments is too small to find a main effect for **environment**.

Whether the latter finding would be significant for phonological theory is a matter for debate. Second, if such perceptually-based adaptive strategies exist, they might manifest themselves more readily in types of speech that are more obviously listener-oriented, such as a conversation in a noisy room or infant-directed speech. Third, if one chose to take the position that deletion is more prevalent in fast speech exclusively because fast speech requires more effort, and not because deletion in fast speech is less noticeable (following Kirchner's (2001) effort-based approach to consonant lenition), then it can be argued that the P-map hypothesis does not predict that deletion should be more frequent in perceptually weak environments in some speech styles and not in others.

While these caveats are duly noted, this study provides experimental evidence that is convergent with a conclusion that has been reached by many phonologists, i.e., that sound change is not part of the grammar (Ohala 1995), that diachronic phonology is not the same as synchronic phonology (Hyman 2001), that phonology does not need to explain all historical facts (Buckley 2000), that phonetic grounding explains facts which can be derived without reference to phonology (Hale and Reiss 2000), and that functional factors such as perception, production, generalization, and conformity can be construed as diachronic filters (Hume and Johnson 2001).

Functional explanation can exist without being in the grammar, so harmonic constraint rankings (which are used in Optimality Theory in an effort to make similar typological predictions) may be epiphenomenal rather than explanatory. Given what is known about the requirements of production, perception, communication, learnability, cognition, social identity, language change, etc., how much crosslinguistic regularity of phonological patterning (phonetically-motivated or otherwise) has a grammar-external explanation? To what extent are other phonological constructs such as distinctive features epiphenomenal rather than explanatory? (Mielke 2003, In prep.)

5. Conclusion

The explanation for why /h/ deletion is more prevalent before and after voiceless fricatives and after voiceless stops involves a conspiracy of perceptual and articulatory factors. Although this phonological process can be explained functionally, no listener-oriented adaptive strategies are apparent in the speech of the talkers in the experiment, indicating that this functional explanation lies in the domain of sound change rather than in synchronic phonology. Therefore it is shown that speakers can possess functionally-motivated phonological rules without any synchronic access to the functional motivation. This conclusion supports removing the burden of explanation from formal synchronic accounts of phonological alternations when explanation can be found in language change and language use. Functional factors certainly explain much about observable patterns in phonological systems, and functional explanation is crucial to explaining how individual phonological systems came to be the way they are, but it does not necessarily explain how phonology works in real time inside the mind of each speaker.

References

Anderson, S. 1981. Why Phonology Isn't "Natural". *Linguistic Inquiry 12*(4): 493-539.

Archangeli, D. and D. Pulleyblank, 1994. *Grounded Phonology.* Cambridge: MIT Press.

Buckley, E. 2000. What should phonology explain? Handout from SUNY Buffalo linguistics colloquium.

Dolbey, A.E. and G.Ó. Hansson (1999) The source of naturalness in synchronic phonology. *CLS 35-1*, 59-69. Chicago: CLS.

Flemming, E.S. 1995. Auditory representations in phonology. PhD dissertation, UCLA.

Hale, M., and C. Reiss. 2000. Phonology as Cognition. In N. Burton-Roberts, P. Carr, and G. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*, 161-184. Oxford University Press.

Harnsberger, J. D., and D. B. Pisoni 1999. Eliciting Speech Reduction in the Laboratory II: Calibrating Cognitive Loads for Individual Talkers, in *Progress Report No. 23, Research on Spoken Language Processing*. Indiana University.

Hayes, B. 1997. Phonetically driven phonology: The role of Optimality Theory and Inductive Grounding. UCLA ms.

Hume, E. and K. Johnson. 2001. A Model for the Interplay of Speech Perception and Phonology. In E. Hume and K. Johnson (eds.) *The Role of Speech Perception in Phonology*, 3-26. New York: Academic Press.

Hyman, L. M. 2001. The Limits of Phonetic Determinism in Phonology: *NC Revisited. In E. Hume and K. Johnson (eds.) *The Role of Speech Perception in Phonology*, 141-185. New York: Academic Press.

Janda, R., and B. Joseph. 2002. On language, change, and language change – or, of history, linguistics, and historical linguistics. In R. Janda and B. Joseph (eds.), The Handbook of Historical Linguistics. Oxford: Blackwell.

Kirchner, R. 2001. *An effort based approach to consonant lenition.* New York: Routledge

Lewis, G. L. 1967. *Turkish Grammar.* Oxford: Clarendon Press.

Lindblom, B. 1990. Explaining phonetic variation: a sketch of H&H theory. In W.J. Hardcastle & A. Marchal (eds.) *Speech production and speech modeling*, 403-440. Dordrecht: Kluwer Academic Publishers.

Mielke, J. 2000. A perceptual account of Turkish [h] deletion. Paper presented at the Montréal-Ottawa-Toronto Workshop on Phonology, Toronto.

Mielke, J. 2002a. Formalizing the perception-phonology interaction: Looking for evidence of the P-map in speech style variation. Paper presented at the Montréal-Ottawa-Toronto Workshop on Phonology, Montréal.

Mielke, J. 2002b. Turkish /h/ deletion: evidence for the interplay of speech perception and phonology. In M. Hirotani, ed. *NELS 32*, 383-402. Amherst, Mass.: GLSA.

Mielke, J. 2003. How to explain natural classes without universal distinctive features. Paper presented at the Montréal-Ottawa-Toronto Workshop on Phonology, Toronto.

Mielke, J. To appear. The interplay of speech perception and phonology: Experimental evidence from Turkish. To appear in *Phonetica.*

Mielke, J. In prep. The Theory of Emergent Distinctive Features. PhD dissertation, OSU.

Munhall, K. & A. Löfqvist. 1992. Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics 20*: 111-126.

Myers, S. 1997. Expressing phonetic naturalness in phonology. In I. Roca (ed.) *Constraints and derivations in phonology*, 125-152. Oxford: Oxford University Press.

Myers, S. 2002. Gaps in factorial typology: The case of voicing in consonant clusters, University of Texas at Austin ms.

Ohala, J. 1981. The listener as a source of sound change. In Masek, C.S., R.A. Hendrik, M.F. Miller, eds., *CLS 17-2*, 178-203. Chicago: CLS.

Ohala, J. 1993. The perceptual basis of some sound patterns. In D.A. Connell and A. Arvaniti, eds., *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*, 87-94. Cambridge: Cambridge University Press.

Ohala, J. 1995. Phonetics and Historical Phonology, UC Berkeley ms.

Pater, J. 1996. *NC. In K. Kusumoto, ed. *NELS 26*. 227-239. Amherst, Mass.: GLSA.

Sezer, E. 1986. An autosegmental analysis of compensatory lengthening in Turkish. in L. Wetzels & E. Sezer, eds. *Studies in Compensatory Lengthening.* Dordrecht. Foris Publications.

Steriade, D. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. ms. UCLA.

Steriade, D. 2001. Directional asymmetries in place assimilation: a perceptual account. In E. Hume and K. Johnson (eds.) *The Role of Speech Perception in Phonology*, 219-250. New York: Academic Press.

Steriade, D. In press. The Phonology of Perceptibility Effects: the P-map and its consequences for constraint organization. In *A Festschrift for Paul Kiparsky.*

Jeff Mielke
222 Oxley Hall
1712 Neil Ave
Columbus, OH 43210

mielke@ling.ohio-state.edu

Accepting Markednesslessness: How non-phonological symbolic computation shapes trends in attested phonological systems

CHARLES REISS
Concordia University

1. Introduction

This paper makes some programmatic suggestions concerning how properties of general auditory processing provide insight into the kinds of sound changes that recur in the transmission of language, thus giving rise to the kinds of recurring patterns seen in the phonological systems of the world's languages. The position developed here must be clearly distinguished from any attempt to build phonetics into phonology, or to blur the distinction between various domains—in fact, one subgoal of this work is to develop coherent and useful definitions of various domains of inquiry, for example, synchronic linguistics *vs.* diachronic linguistics (two fields of study, one of which examines the human language faculty) and auditory perception *vs.* phonology (both of which examine faculties of the mind, one that interfaces with the language faculty, and one that is a component of the language faculty).

I hope to show that rejecting substantive 'markedness' (cf. Hale & Reiss 2000ab) as a crucial part of phonological theorizing does not force us into theoretical nihilism. There is still plenty of work to do in 'psychological phonology', and there is still plenty to do in related fields, even if we accept 'markednesslessness'.

Drawing on recent work in auditory perception in the auditory scene analysis framework, I hope to support the research program founded by John Ohala which aims to explain phonological patterns by appeal to extragrammatical phenomena, as alluded to in this recent quotation:

> So, the phonetic primitives invoked in the modeling of these [diachronic] processes make no pretenses of being psychological. The attempts by

> those who are interested in psychological phonological grammars and in finding ways to represent phonological processes ... in phonetically natural ways have been abysmal failures ... One possible solution to this is not to put more phonetic sophistication into psychological grammars but rather to abandon phonetic naturalness as a necessary feature of them.
>
> Ohala 2002:685

This approach is to be contrasted with the currently dominant trend within linguistics of building more and more phonetics into phonology, for example, in the form of phonetically grounded grammatical constraints.

2. Background assumptions and terminology

Following Keating (1988) I adopt the view that phonology is most usefully defined to consist of a particular kind of symbolic computation. The input and output of the phonology are representations described in terms of a single symbolic alphabet. The output of the phonology is *transduced* to actual articulatory and perceptual events *via* a series of transducers, some of which are also cognitive, but by definition non-phonological. This modular approach to explanation is consistent with standard views within cognitive science:

> In using the word 'representations', we are implying the existence of a two-part system: one part forms the representations and another uses them to do such things as calculate ...
>
> Albert Bregman, *Auditory Scene Analysis* (1990:3)

It is crucial to note that cognition builds and manipulates EQUIVALENCE CLASSES—it is not the case that every physically distinguishable contrast is computationally/psychologically relevant (cf. Pylyshyn 1984, and even Sapir 1933).

In order to contrast the focus of this paper with what I take to be useful markednessless phonological theorizing, I will mention two examples of the latter. Work by Sharon Inkelas (*e.g.*, 1996) on the form of featural representation in the lexicon, including issues such as the justification for underspecification constitutes valid, interesting phonological work. My own work attempting to demonstrate the necessity of the use of quantifiers in phonological computation and the implications of this necessity, *viz.*, that feature geometry is *not* a property of phonological representations is another example of work that falls squarely within the tradition of phonological reasoning in a generative framework. In the next section we turn from phonology to a discussion of auditory perception.

3. Auditory Scene Analysis

> The perceptual world is one of events with clearly defined beginnings and endings ... An event becomes defined by its temporal boundary. But this impression is not due to the structure of the acoustic wave; the beginning and ending often are not physically marked by actual silent intervals.
>
> Handel, 1989

Auditory scene analysis is a framework for studying auditory perception developed by Albert Bregman and his collaborators. I think Bregman would be the first to admit that work in the field is still in its infancy. However, it has now become possible to ask questions concerning the nature of auditory perception that approach the sophistication of question in domains such as visual perception.

Auditory scene analysis can be broken down into two main components. One problem, given the fact that sounds waves from various sources are combined into a single wave that reaches the eardrum, is that of simultaneous integration and segregation. In other words, the auditory system integrates into a single representation parts of the sound spectrum reaching the ear within a temporal window that 'go together', according to the properties of the system (nonveridically, in the case of an illusion). The process of assigning parts of the spectrum to different perceptual sources is called spectral segregation.

The other main component of auditory sene analysis is sequential integration—acoustic events occurring separated in time may be integrated into a single auditory stream. Examples of streams include a sequence of footsteps or the continuous sound of falling rain.

3.1. A Simple Grammar for auditory scene analysis

Building on the work of Y. Nakajima and T. Sasaki, Nakajima (1996) proposes that auditory scene analysis can best be understood in terms of a grammatical system that recognizes elements smaller than the auditory events (a footstep, a musical note, *etc.*) that make up auditory streams. These elements, auditory subevents, are classified into at least four fundamental types.

(1) Types of auditory subevents

- Onset (denoted by <): A steep rise of sound intensity within a certain frequency range (e.g., a critical band) can be a clue of an onset.
- Termination (denoted by >): A steep fall of sound intensity within a certain frequency range can be a clue of a termination.

- Filling (denoted by =): A piece of sound energy extending for a certain duration without any sudden change of frequency range can be a clue of a filling.
- Silence (denoted by /): If the sound energy across a certain frequency range and a certain duration is very thin despite some amount of sound energy in the preceding part, this makes a clue of a silence.

Onsets and terminations will be referred to below as boundaries.

As these definitions make clear, what is of interest to us are the cognitive representations transduced from the physical signals. So, for example, the auditory grammar subevent 'silence' can be transduced in a context of *acoustic* energy, for example, where the energy is not part of the same stream as the perceptual 'silence'. It is sometimes difficult in discussion to be consistent in distinguishing the cognitive subevents from their acoustic correlates. However, use of the symbols mentioned above will be helpful: the symbol '/' is a constructed representation that may be transduced during a period of acoustic energy, for example.

According to Nakajima, most auditory events in our everyday life seem to take one of the three modes indicated below.

(2) Auditory events:

a. An onset followed by a silence (e.g., a clap sound): (<)/ ,
b. An onset, a filling and a termination followed by a silence (e.g., a cat's meow): (<=>)/ ,
c. An onset and a filling followed by the onset of another auditory event (e.g., a note in a melody played legato): (<=)< .

Following Nakajima, the auditory events we find are represented by the portions within parentheses. The subevents, following the right parenthesis represent the beginning of the next auditory event or a silence.

The primitive auditory subevents are thus grouped into events which are the immediate constituents of auditory streams. The events are combined into streams acording to a small set of principles.

(3) Definition of auditory stream:

a. An auditory stream is a linear string of auditory events and silences in temporal order.
b. An auditory stream begins with an onset and ends with a silence.
c. A silence is not followed immediately by another silence.

To summarize the model developed thus far, an auditory scene is composed of auditory streams that are composed of auditory events that are composed of auditory subevents.

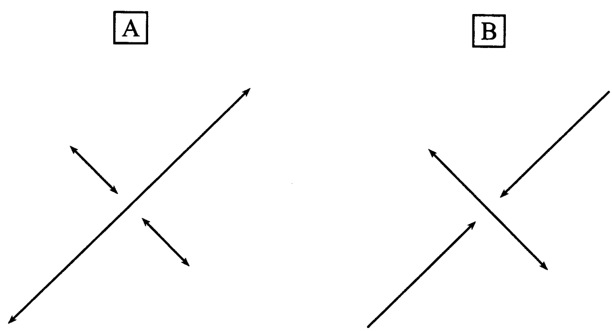
4. Gap transfer illusion

In this section we discuss an auditory illusion, the gap transfer illusion, discovered by Nakajima and Takayuki Sasaki that suggests that "When it is difficult to interpret all the given clues of auditory subevents grammatically, the auditory system inserts new subevents, interprets the same clues twice or more, or suppresses some clues." It is crucial to keep in mind that the subevents are discrete symbols over which auditory scene analysis computes. The four primitives are manipulated by the computational system in the construction of the auditory scene. Gestalt principles such as proximity and similarity can explain some of the manipulation of subevents by cognition.

Our discussion of the gap transfer illusion will necessarily be rather informal. In the diagrams below, arrowheads are to be equated with the boundary symbols, < *and* >, introduced above. Lines correspond to fill (=). The *x*-axis corresponds to time and the *y*-axis to frequency.

The gap transfer illusion arises when a stimulus such as (4B) is perceived as (4A).

(4) Gap transfer illusion



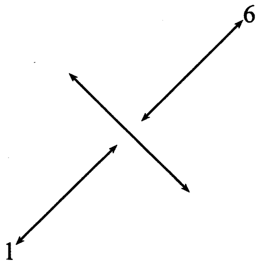
B perceived as A.

That is, the gap in the acoustic rising glide is transferred to the falling glide.

There are actually several things going on here, some of which are strictly speaking not part of the gap transfer effect. It will be useful to discuss them, however, to illustrate the working of the auditory grammar that Nakajima and his colleagues propose. First, note that the stimulus in (4B) does not begin with an onset, <, or end with a termination, >. This is because the rise and fall of the sound intensity in these locations were not 'steep enough to give such clues', according to Nakajima. A filling at the beginning of an auditory stream or a filling immediately followed by a silence is ungrammatical. That is, such contexts cannot be derived from the above-mentioned grammatical rules. Apparently, the solution for the auditory system is to insert an onset (in accordance with (3b)) and a termination (in accordance with (2b)).

The insertions demanded by these auditory rules would give something like (5). (Note that I am fudging the issue of what exactly (5) is. It is not exactly a representation of the stimulus and it is not the final percept. Perhaps it is best thought of as an intermediate auditory perception.)

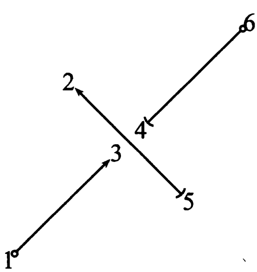
(5) Boundary insertion



The relevant onset and termination are inserted by the auditory grammar at the points labelled '1' and '6'.

We are now ready to describe the Gap Transfer Illusion *via* the Gestalt Principle of Proximity. In order to do this, I have replaced some of the arrowhead onset and termination markers with other symbols. The recoupling or reassociation of these boundary markers can be deduced by matching pairs of symbols. The discussion will also be facilitated by the numerals, corresponding to the physical temporal sequencing, that I have assigned to the boundary markers.

(6) Boundary reassociation



In the signal, or, more precisely, the intermediate representation of (5), boundaries 1 and 3 belong together, as the subscript *a* denotes below in (7). Boundaries 2 and

5, with subscript *b* also belong together. Finally, 4 and 6, with subscript *c*, belong together.

(7) SIGNAL (plus perceptual insertion of 1 and 6)

$$<_{a1} \qquad <_{b2} \quad >_{a3} \qquad <_{c4} \quad >_{b5} \qquad >_{c6}$$

However, the percept is very different. The temporally close onset-termination pair (2,3) have been coupled, as indicated by the shared arrowheads in (6) and by the subscript β that they share in (7). Similarly, the pair (4,5) share the curved boundary markers in (6) and the subscript γ in (7), denoting the fact that they have been perceptually coupled. This leaves the pair (1,6), with open circles for boundary markers in (6) and subscript α in (7).

(8) PERCEPT (after reassociation of boundaries)

$$<_{\alpha 1} \qquad <_{\beta 2} \quad >_{\beta 3} \qquad <_{\gamma 4} \quad >_{\gamma 5} \qquad >_{\alpha 6}$$

The percept resulting from insertion and reassociation of boundaries from an input stimulus like (4B), is thus identical to that resulting from an input stimulus like (4A).

Before moving on to speech perception, let's make two observations. First, the percepts of onsets and terminations, as well as their mutual associations, are *constructed* in the process of auditory stream analysis.

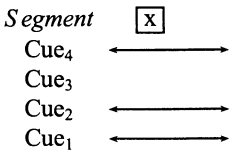
Second, it is suggestive that the effect of reassociation in the gap transfer illusion is to provide the auditory event with an immediate constituent analysis. In the final percept, as denoted in (8), constituent α contains constituents β and γ, and β precedes γ, but there are no interlocking constituents. On a highly speculative note, I suggest that this is the defining characteristic of auditory streams, as well as of linguistic representations. In fact, the perception of speech as a sequence of discrete segments may be a reflection of this property of auditory streams—in the streaming of the diverse components of the speech signal only precedence and containment relations among cues are licit.

5. Application to speech perception

In this section, we treat speech signals in isolation as auditory scenes composed of various auditory streams that correspond to the acoustic cues generated by a human vocal tract in speech. In other words, acoustic parameters such as the value of the first formant or the presence of high frequency broadband noise, each constitute, by hypothesis, an auditory stream. These streams, of course, can be further analyzed into their component events and subsevents.

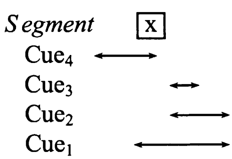
We can imagine an idealized representation of the relationship between the acoustic cues that are transduced to phonological featural representations with perfect temporal alignment for all the cues. In the example in (9) Cue_3 is absent from segment x.

(9) Idealized segment



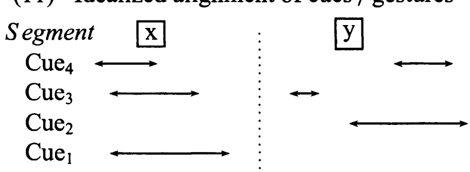
However, it is well known that from either an articulatory or an acoustic perspective, temporal relations of gestures or perceptual cues are much less orderly, as shown in (10), where Cue_1 extends over Cue_3 completely, and Cue_1 partially overlaps with Cue_4, and the latter completely precedes Cue_1.

(10) Acoustic cue / articulatory gesture alignment



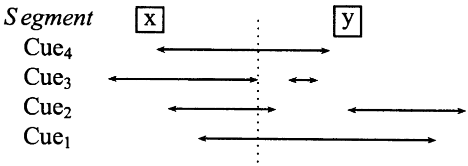
Even granting that cues do not all line up perfectly in duration, we could imagine an idealized representation of the temporal sequencing of cues in a sequence of segments, with a neat division between the cues belonging to two segments x and y, as in (11):

(11) Idealized alignment of cues / gestures



However, once again, the acoustic and articulatory reality is much messier. The alignment relations in (12) are much more realistic.

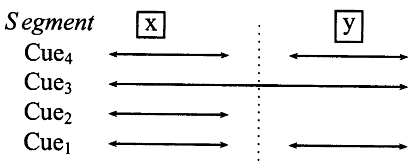
(12) Physical alignment of cues / gestures



In other words, cues that end up being interpreted as belonging to a give segment may partially or even fully overlap temporally with the cues of another segment.

Despite the complex temporal relation among cues both within and across segment boundaries, I would like to suggest that the equivalence classes generated in the process of speech perception lead to a representation more like (13), where I have equated (simplistically) cues with features. In other words, cues are transduced to feature bundles (segments) which are discrete categorical symbols.

(13) Claim: (Quasi-)Phonological representation



We can now proceed to a discussion of how auditory scene analysis and Nakajima's grammar of auditory events can give insight into the explanation for the recurring patterns observed in phonological systems.

5.1. Why are long cues subject to change?

In various works John Ohala and those influenced by him (*e.g.*, Blevins 2003, Blevins and Garrett 2000) have noted that phonological features corresponding to phonetic cues with long durations are more likely to undergo common sound changes such as assimilation, dissimilation and metathesis. Examples cited by Blevins & Garrett include rhoticity (the acoustic correlate of which is lowered F3), palatalization (raised F2), and rounding (lowering of all formants). I would like to make the simple suggestion that the explanation for this behavior is that the acoustic signals corresponding to streams for long duration cues will overlap with more subevents belonging to other streams than the acoustic signals of streams for short duration cues. Thus, there are more possibilities that the boundaries of long duration cues be reassociated by a perceiver with acoustic subevents belonging to other streams in the target language.

Consider first how the Proximity Principle could be invoked to account for assimilation. In (14.i) we represent abstractly two cues/features *p* and *q*, that are attributes of separate segments in the target language. The vertical line is meant to denote a segment boundary.

Now suppose that (14.ii) represented the temporal relations of the boundaries (onsets and terminations) of these cues as produced by speakers of the target language. Cue p corresponds to constituent a and cue q corresponds to constituent b, and there is a partial overlap of p and q. Under appropriate conditions (to be determined by future researchers), the proximity of onset $<_2$ and termination $>_3$ leads to association of these two subevents by the learner/perceiver. If we assume that boundaries $<_1$ and $>_4$ are also coupled, we get a model of the learner's auditory percept that has the structure of (14.iii), where constituent α contains constituent β. Step (14.iv) shows how this percept is transduced to a grammatical representation.

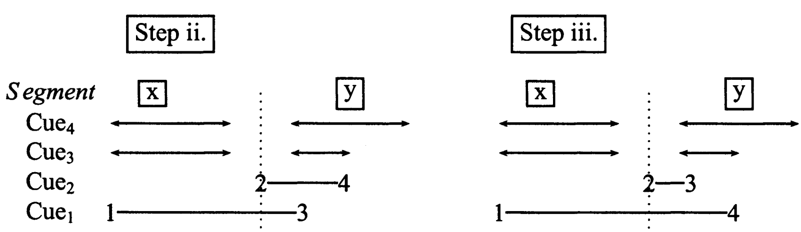
(14) Assimilation (*via* PROXIMITY-driven reassociation)

i. FEATURAL REPRESENTATION IN TARGET LANGUAGE			$[p] \mid [q]$				
ii. CUE PATTERN IN TARGET OUTPUT	$<_{a1}$		$<_{b2}$	$>_{a3}$		$>_{b4}$	
iii. LEARNER'S AUDITORY PERCEPT		$<_{\alpha 1}$		$<_{\beta 2}$	$>_{\beta 3}$		$>_{\alpha 4}$
iv. LEARNER'S GRAMMATICAL (FEATURAL) PARSE			$[p] \mid [p, q]$				

In this grammatical representation, the first segment is still associated with the cue/feature p, and the second with q, but the second is also associated with p. If the learner stores the representation thus, diachronic assimilation has occurred.

Steps (ii) and (iii) can also be represented as follows, with p corresponding to cue 1, and q to cue 2.

(15) Temporal relations of cues



Parallel to the case of the gap transfer illusion, the association of 2 and 3 allows 1 to associate with 4.

It is not difficult to extend this form of reasoning to derive dissimilation of a long duration cue. In the following diagrams, the target language has a segment associated with cue p followed by one associated with p and q. The auditory event corresponding to q either has no termination, or else the termination is masked by other acoustic properties of the stimulus. The dissimilatory change is the loss of association of p with the second segment.

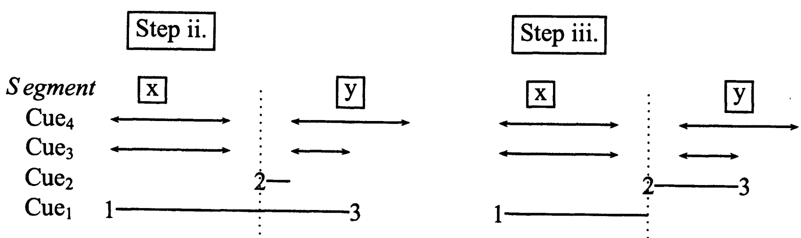
(16) Dissimilation (*via* PROXIMITY-driven reassociation)

i. FEATURAL REPRESENTATION IN TARGET LANGUAGE $[p] \mid [p, q]$

ii. CUE PATTERN IN TARGET OUTPUT $<_{a1}$ $<_{b2}$ $>_{a3}$

iii. LEARNER'S AUDITORY PERCEPT $<_{\alpha 1}$ $<_{\beta 2}$ $>_{\beta 3}$

iv. LEARNER'S GRAMMATICAL (FEATURAL) PARSE $[\alpha] \mid [\beta]$



The association of 2 and 3 leads to the creation from 1 of an event of type (2c), 'an onset and a filling followed by the onset of another auditory event', so Cue_1 does not extend through the second segment in the learner's parse.

6. The CRUM-my Conclusion

Obviously much more work is needed to make the suggestions here more concrete. The application of the Principle of Proximity and its interaction with other Gestalt principles is poorly understood, and it will be necessary to look at real examples of assimilation and dissimilation to see if Nakajima's auditory grammar will work for a signal as complex as speech. However, the notion of unifying aspects of phonology with aspects of speech perception and ultimately with aspects of auditory perception is appealing and warrants further research. Note that this unification does not entail a rejection of phonology as a domain of inquiry—aspects of phonological systems (and phonological typology) that can be explained by reference to audition do not have to be explained by reference to the faculty of phonological computation. This leads to a more elegant theory of phonology.

I adopt the position of classical cognitive science that cognition can best be understood in terms of representational structures composed of discrete, categorical primitives and computational procedures that operate on those structures: this view is dubbed the the Computational-Representational Understanding of Mind (CRUM) by Thagard (1996). Generative phonological theory is an example of a CRUM-my theory. (That is a good thing.)

The thesis of this paper is that the recurring patterns (assimilations, dissimilations, *etc.*), that we find in phonological systems are due to a CRUM-my module of mind. However, those patterns are the result of the CRUM-my nature of auditory scene analysis applied to speech during acquisition, and not the computational properties of the phonological component. This position is consistent with the diachronic, nonfunctionalist approach to sound change and phonological typology represented by the work of Ohala, Blevins & Garrett and Hale.

I conclude then, that it is not necessary to build the details of phonetic substance into the phonological component by encoding 'markedness' based on typological tendencies into grammar. There is plenty of real phonology to be done, and there is plenty of work in fields such as auditory perception and speech perception. In other words, a markednesslessnessless phonology is not needed to combat nihilism.

One final observation should be made. If we accept the analysis of the gap transfer illusion presented by Nakajima and his colleagues, then we accept the possibility that the auditory system manipulates entities like terminus markers in a way that is highly abstracted from their physical substance, however that may be defined. It is no surprise therefore that phonology, which in turn is several steps removed from auditory processing should also show evidence of substance-free computation.

References

Blevins, Juliette. To appear. *Evolutionary Phonology*. Cambridge University Press.

Blevins, Juliette and Andrew Garrett. To appear. The evolution of metathesis. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.). *Phonetically driven phonology*. Cambridge University Press.

Blevins, Juliette and Andrew Garrett. 1998. The origins of consonant-vowel metathesis. *Language* 74: 508-56.

Bregman, A. 1990. *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Hale, M. To appear. *Theory and Method in Historical Linguistics*. Blackwell.

Handel, S. 1989. *Listening*. Cambridge, MA: MIT Press.

Inkelas, Sharon. 1996. Archiphonemic Underspecification. Ms. UC Berkeley.

Nakajima, Y. 1996. A simple grammar for auditory organization: streams, events and subevents. *Approaches to Auditory Organization" XXVI International Congress of Psychology*, Montreal, Canada

Nakajima, Y., Sasaki, T., Kanafuka, K., Miyamoto, A., Remijn, G., and ten Hoopen, G. (2000). Illusory recouplings of onsets and terminations of glide tone components. *Perception & psychophysics, 62*, 1413- 1425.

Pylyshyn, Z. 1984. *Computation and cognition: toward a foundation for cognitive science.* MIT Press, Cambridge, MA.

Reiss, Charles. 2003. Quantification in phonological rules. *The Linguistic Review 20: Special Issue on Typology in Phonology*, Spring 2003.

Sapir, E. 1933. La Réalité Psychologique des Phonèmes. Journal de Psychologie Normale et Pathologique (Paris), 30: 247-265.

Knowledge of similarity and narrow lexical override

DONCA STERIADE
Massachusetts Institute of Technology

0. Introduction

This study summarizes the findings of an investigation into patterns of half-rhyme use in mid-20th century Romanian poetry. The data suggest that linguistically naïve speakers – here Romanian poets and their readers - possess detailed, shared knowledge of context-sensitive relations of phonological similarity. I show that knowledge of sound similarity cannot be deduced from knowledge of lexical patterns (i.e. sound distributions or patterns of alternation). I propose that it is based on speakers' understanding that the distinctiveness of contrasts – the effectiveness of the cues that differentiate a contrasting pair - varies with context. Speakers know that pairs of contrasting sounds are more similar if placed in a context where their differentiating cues are less effective.

1. Half-rhymes

Half-rhymes (HR) are rhyming pairs that fall short of strict identity. Thus Bob Dylan rhymes *ez* with *ezd* in *'all my days/... partly raised* '(Golston 1995). Some HRs are judged closer to identity than others: we infer this from the fact that some HRs are more frequent than others. In the rock-poetry studied by Zwicky (1976) and Golston (1995) HRs like *ez-ezd* are about as frequent as perfect feminine rhymes, and much more frequent than HRs like *'all my days/...all the pretty maids',* in which *ez* is matched with *edz*.

There is a judgment of relative similarity implicit in the frequency difference. I use the notation in (1) to summarize such judgments:

(1) $\Delta(x\text{-}y) > \Delta(z\text{-}w)$.

This means "the perceived difference between the strings x and y is greater than that between strings z and w", hence x-y is less similar than z-w. The relative similarity judgment discussed above corresponds, in this notation, to (2):

(2) $\Delta(\text{ezedz}) > \Delta(\text{ez-ezd})$.

The better HR is the more similar pair *ez-ezd*. We can generalize (2) as follows:

(3) $\Delta(\text{C-}\emptyset)//\text{V} > \Delta(\text{C-}\emptyset) \neg(//\text{V}))$

(3) says that the difference between a string containing a C and an otherwise identical string lacking this C is greater when the C is V-adjacent (i.e. in the context //V) than otherwise (in the context ¬(//V)). This similarity judgment – like that in (2) - is context-dependent: the same C-Ø contrast counts for more in some contexts than in others. Existing models of segmental similarity (Frisch, Broe, Pierrehumbert 1997) operate in context-free fashion and cannot characterize this fact: they are thus unsuited to model the similarity knowledge underlying HR use.

Our hypothesis is that the computation of similarity is context-sensitive because the judgment of similarity is, in part, a context-dependent judgment of relative perceptibility. The transitions into and out of a vowel provide cues to the nature of a consonant's oral constriction. The difference between a V-adjacent C (the [d] of *edz*) and a non-V adjacent C (the [d] of *ezd*) is then one of relative perceptibility: the V-adjacent [d] is more perceptible and thus its absence (in the *edz-ez* HR) is more salient. The non-V adjacent [d] yields the better *ez-ezd* HR because it is less perceptible.

Poets decide which HRs to use; their audience decides which HRs are tolerable. Our experience as readers suggests that these two judgments converge: the more frequently employed HRs are also the ones that readers perceive as more similar. The idea of a widely shared set of relative similarity judgments can be verified by observing if the relative frequencies among HR types are similar in different poetic corpora. The idea that many similarity judgments are shared by poets is tacitly assumed by earlier writers on the subject[1], who pool the HR data of multiple poets rather than separately studying each corpus; this idea is documented below.

HR preferences merit investigation because the shared knowledge of relative similarity they reveal is, in Halle's (1978) terms, "knowledge untaught and unlearned". In this they differ from the knowledge speakers have of the patterns of alternation and sound distribution of their language: these are learned by learning words. Learning one's lexicon doesn't, by itself, teach us which pairs of words are more similar. In most poetic traditions there are no publicly posted standards of HR use. Poets are taught by example what counts as a perfect rhyme, but they're on their own when they need to find the minimal deviation from rhyming identity: it is then remarkable if the HR choices of several poets converge on a shared hierarchy of similarity. How is this knowledge achieved? And how does it relate

[11] Eekman (1974), Zwicky (1976) and Holtman (1996)

to phonological computations? I defer a partial answer to the first question to section 10. The second question is taken up below.

2. Uses of similarity: the P-map hypothesis

Elsewhere (2001, 2003) I have suggested that certain grammatical properties are inferred by learners from their knowledge of relative similarity relations (referred to as "the P-map"). The hypothesis is that learners induce rankings among correspondence constraints from known rankings of similarity: if a learner knows that Δ(C-Ø)//V > Δ(C-Ø) ¬(//V)) he infers that any correspondence condition prohibiting a C//V vs. Ø difference between some pair of representations (say input and output) outranks the correspondence condition that prohibits the less salient C¬(//V)) vs. Ø difference between *that* pair of representations (i.e. input and output). This amounts to saying that knowledge of the similarity relation in (3) causes the learner to infer the ranking in (4):

(4) MAX C//V (IO) >> MAX C¬(//V)) (IO)

The HR data discussed here bears on several aspects of the P-map hypothesis. The evidence suggests specific links between context-dependent differences in cue distribution and context-dependent similarity judgments. It also allows us to explore possible relations between lexically manifest patterns of sounds and knowledge of similarity (cf. Frisch, Brie and Pierrehumbert 1997). In particular we consider here cases in which knowledge of perceptual similarity projects correspondence rankings that differ from those supported by the lexical evidence.

A more direct answer to the question "Is knowledge of similarity relevant to phonology" arises in the study of loan adaptation. This activity also relies on untaught judgments of relative similarity. For any borrowed word in need of nativization, there are multiple adaptation strategies. Choices among these rely on the same similarity computation as those that guide poets in the search of the better HR. Consider for instance a Hausa speaker who borrows English *yeast* as [jis] (Newman 2000). This borrower has to decide whether to simplify the English C_1C_2 coda via epenthesis or via deletion of C_1 or of C_2. The phonology of Hausa has no relevant alternations to break the tie in favor of one of these strategies; and a speaker who can't identify other comparable loans into Hausa does not know how others have solved this problem. Then he is on his own in this choice, exactly like the poet looking for the better HR. A further parallel with the rhyming case is that the choice made in borrowing a CVC_1C_2 syllable (delete the non-V adjacent C_2, not the postvocalic C_1) reflects the ranking of correspondence in (4), and thus the similarity ranking in (3). More generally, the study of coda cluster simplification in loan adaptation supports the idea that the ranking in (4) is constant across different borrowers (Silverman 1992, Steriade 2001), even if their native language lacks informative alternations. The P-map hypothesis offers the same answer to the borrowing question (How does the borrower choose which C delete in CVC_1C_2?) as to the rhyming question (How does the poet know which C is easier to miss: the post-vocalic C_1 or the non-postvocalic C_2?). The answer is in both cases: the least perceptible C. The suggestion, to repeat, is that knowledge of context-dependent perceptibility translates into rankings of similarity which in turn yield correspondence rankings. The latter operate in loan adaptation; they arguably function in the analysis of rhyming as well (Steriade and Zhang 2001).

3. The Romanian corpus

The collection of Romanian poetry analyzed here consists of six rhymed translations and three native poetry texts. The translations are, with one exception, from Russian, and they were all published between 1956-1971: see the appendix for details. The entire translation corpus totals 9791 rhyming pairs, of which 693 pairs are HRs. HR frequency in the six translations range from a high of 18% (in the translation of Briusov poems by Kernbach) to a low of .02% (in the translation of Pushkin by Lesnea.)

The native poets studied are Tudor Arghezi (1956), Miron Radu Paraschivescu (1957) and Niculae Labis (1959). There are 9223 rhyming pairs to 323 HRs in these texts. The HR frequencies are 0.58% (Arghezi), 10% (Paraschivescu) and 4.7% (Labis). The entire poetry-and-translation corpus is linguistically homogeneous: the poets write in Standard Literary Romanian (SLR); with the exception of a few of Arghezi's early poems, the texts were composed from the mid 40's to the late 60's; and, with the possible exception of Niculae Labis, the poets also *spoke* SLR, not a rural dialect.

Romanian poets differ in their tolerance for HRs. The poets I discuss use significant quantities of HRs, compared to others, but even for them the HRs are a very small fraction of overall rhyming pairs: this suggests that the poets distinguish full from partial identity and only *tolerate* the latter.

The data was analyzed to answer three questions. First, do the poets share a hierarchy of similarity, as evidenced by their common preferences for some HRs. If they do, is this hierarchy related to perceptibility factors: are feature mismatches more common in contexts where the relevant contrasts are less distinctive? Finally, do lexical patterns of alternation and distribution influence the poets' judgment of similarity?

4. Shared knowledge of similarity

On the issue of the shared similarity hierarchy, the first relevant observation is that there is a core of HR types that most poets will employ, and then there is a periphery whose use is sporadic both within and across different authors. A summary of the core types appears in (5). The rightmost column provides the ratio of HRs of the type identified in a given row to the other HRs found in a given text. The texts are identified by the first initial of the author (with key provided in the Appendix): thus the S37% figure in the rightmost top cell states that postnasal voicing - mismatches like *pʌmĩnt -strĩŋgĩnd* 'earth-squeezing' represent 37% of the HRs found in Petre Solomon's Shelley translation.

(5) Core HR types

Mismatched contrast	**Example**	**Frequency: this HR type/all HRs in a given corpus**
Voice/[+nas]_	*pʌmĩnt - strĩŋgĩnd*	S37%, B27%, L31%, A25%, P16% M10%, D9%,
Voice/[-nas]_{#, C}	*pantof – popov*	A17%, B14%, M10%, L 11%, D4%, S3%
j-Ø/C_#	*azj – obraz*	S36%, D27%, B24%, L7%, P7%, M6%
(ɨ-i)// [+nasal]	*strĩmte - sĩmte*	D13%, P 13%, L6%, M5%, S3%
(u-ɨ)// [+nasal]	*sũnt - pʌmĩnt*	
Liquid-Ø/Back V_C	*surd – aud*	P16%, M6%, D 6%, A4%,
Liquid-Ø/ CC_	*adʌpostu– nostru*	M7%, L1%, S 3%
m-n/_{#, C}	*strĩmt – vĩnt*	L 4%, M3%, A3%, S3%

Two observations were verified against a larger selection of Romanian poetry (native and translation). First, every poet that permits HRs (and for some poets the total may be as low as 8 HRs), will draw mostly on the core types in (5). Second, these core HRs look like familiar phonological processes (postnasal voicing, final devoicing, depalatalization, nasal vowel centralization, liquid loss, nasal place neutralization) but these are not processes active in Romanian. There are no neutralizations of these contrasts in these contexts in Romanian, and there are no relevant alternations.

5. Voicing HRs

We explore now in greater detail the idea of a shared similarity hierarchy. The most commonly mismatched feature is voicing and the most common contexts where it is mismatched are post-nasally and word-finally. Figure 1 shows that every poet in the corpus analyzed has more voicing HRs postnasally than in the post-oral context _#. This is the first clear evidence of a shared hierarchy of similarity for all poets: all show greater use of one HR type over the other.

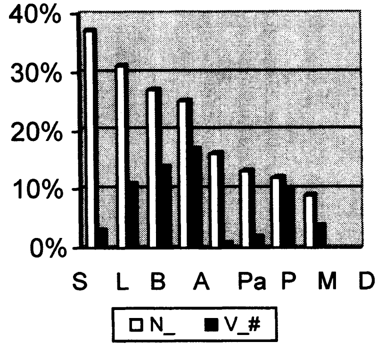


Figure 1: Voicing mismatches in two contexts:
• % = voice mismatch/all HRs in corpus.
• Corpora identified by initials: see appendix.

However, different sequences occur with different frequency in the rhyme domain: specifically, we must control for the possibility that voiced stops are more frequent post-nasally. That alone could explain the data. We compare then four of the corpora with respect to the following ratios: voicing HRs in a specific context/out of all rhymes containing obstruents in the rhyming domain in that context. Thus, if a context is infrequent overall or infrequent in a given corpus, this procedure will factor out this effect. I considered four contexts: V_V (e.g. *fáta-láda*), V_# (*nás-extáz*), N_V (*únde-múnte*) and N_# (*tímp-skímb*). Figures (2)-(5) provide information on the relative frequency of voicing HRs in these contexts for four corpora (two translations and two original poetry collections):

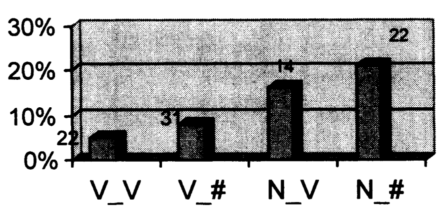


Figure 2: Voicing mismatches in Maiakovsky (Teodorescu)

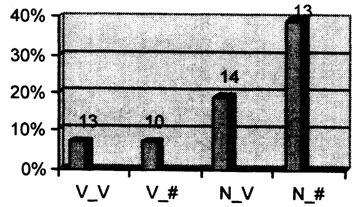


Figure 3: Voicing mismatches in Briusov (Kernbach)

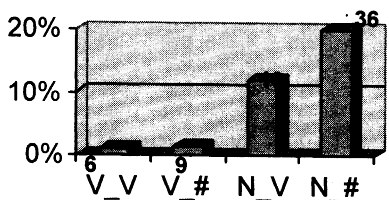


Figure 4: Voicing mismatches by context in Labis

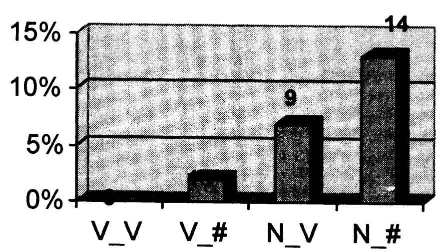


Figure 5: Voicing mismatches in Arghezi

For all four poets, there are significantly more voicing HRs in the post-N than post-V coda and significantly more voicing HRs in post-N codas than in post-N onsets. The differences between intervocalic and postvocalic HRs are not significant but a trend is constant across 3 corpora: there are more mismatches in the coda position. Thus, although the rhymes differ, the frequency of voicing HRs differs, and the frequency of the contexts in which the rhyming obstruents occur differs, the context-sensitive hierarchy of preference for voicing HRs is identical. This shared hierarchy is outlined in (6):

(6) $\Delta(\pm\text{voice}/V\_V) \geq \Delta(\pm\text{voice}/V\_\#) > \Delta(\pm\text{voice}/N\_V) > \Delta(\pm\text{voice}/N\_\#)$

Where does knowledge of this hierarchy come from? In a study of two Romanian speakers, Steriade and Zhang (2001) measured the voicing ratio (the duration of the voiced interval divided by the duration of the voiced+voiceless interval in a consonantal sequence) of voiced and voiceless stops (*b, d, g; p, t, k*) in six contexts: after three nasals (*m, n, ŋ*) and three vowels (*a, i, u*). The difference in voicing ratios of voiced and voiceless stops, in a given context, is an indication of their acoustic difference in that context. One result was that for both speakers, the difference in voicing ratios was significantly smaller after nasals than after vowels: in other words, one dimension of the acoustic distance between voiced and voiceless stops – the duration of the voiceless interval - is reduced after nasals. This is the effect of gradient post-nasal voicing (Hayes and Stivers 1996). The preference for postnasal voicing HRs found in Romanian poetry suggests that poets are aware of – and exploit - the perceptual effects of this phonetic process.

The same study found that the voiced-voiceless contrast is invariably realized as a VOT difference (consistently negative for voiced stops; up to 25 ms VOT for [p], [t]; and up to 55 ms VOT for [k]) in prevocalic position, whether post-nasal or not. The voicing contrast is also expressed in word-final releases, but variably: there are multiple realizations for both voiceless and voiced stops in that context. Thus the lack of consistent release-related cues to voicing may render the final contrast less discriminable. This is consistent with the finding that there are slightly more voicing HRs in post-vocalic final position than prevocalically; and with the very significant increase in HR frequency in the N_# context compared to the N_V context, observed in figures 2-5. Here too the poets' use of final voicing HRs reflects a context-sensitive, phonetically-informed similarity computation.

6. Backness HRs

All but one of the poets studied tolerate vowel quality mismatches. The most common vocalic HRs are the backness mismatches involving nasalized ũ-ɨ̃, ĩ-ɨ̃ pairs in pre-NC position: e.g. *karɨ̃mb-plũmb* 'sole-lead' and *sɨ̃mt-strɨ̃mt* 'I feel-narrow'.

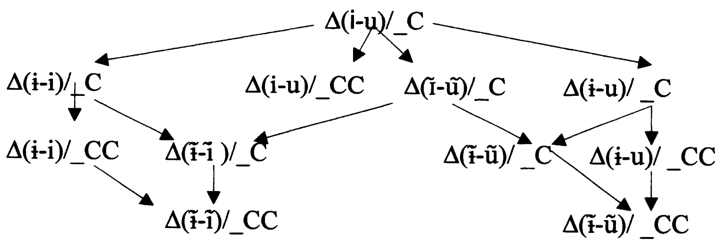
Romanian vowels are nasalized before nasals (Avram 1968); they are shorter before CC clusters. Nasalized *ɨ̃* and *ũ* are separated by a smaller F2 distance relative to oral *i-u* (Wright 1986), a result that carries over to Romanian. F2/F3 values for Romanian oral high vowels appear in (7).

(7) F2/F3 values for Bucharest Romanian oral *i, ɨ, u* (Suteu 1963)

	u	ɨ	i
F2	800	1500	2200
F3	2500		2900

The *u-ɨ, ɨ-i* distances shown above will be reduced by nasalization. If the distinctiveness of vocalic contrasts is in part a function of vowel duration, then the front/back distinctions in the shorter vowels before CC will be diminished relative to the longer vowels of the same quality that precede a single C. These three considerations (greater distance in F2 between *u-i* vs. *u-ɨ, ɨ-i*; the effect of short duration on vocalic distinctiveness; the effect of nasalization on distinctiveness of F2-based contrasts) combine to yield the following expected set of perceptually based similarity relations:

(8)



(8) identifies as least distinctive the contrasts between shortened *ɨ̃-ĩ* and *ɨ̃-ũ* before NC. The HR evidence identifies these as the most similar pairs.

A subset of the poets who admit HRs like *karɨ̃mb-plũmb* also rhyme *ɨ̃-ĩ* and *ɨ̃-ũ* before single nasals, as in *sũnʌ-tsʌrɨ̃nʌ* 'sounds-soil'. These HRs are much less frequent, as shown in figure (6), and as predicted by (8).

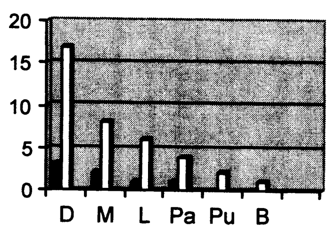


Figure 6: frequency of *ɨ̃-ĩ* and HRs *ɨ̃-ũ* by context:
Light columns: before _NC. Dark columns: before N.

A further subgroup of poets rhyme oral *ɨ-i* and *ɨ-u*, in specific contexts (_j, //r), as in *ɨ̃ntij-gutuj* 'first; quinces'. These cases are too rare to yield significant counts in each corpus but the contexts are significant, as both [j] and [r] influence vowel quality and thus are likely to diminish the perceptibility of the backness contrast. Only the most HR-tolerant poets (such as Maiakovsky's translator Cicerone Teodorescu) admit, rarely, oral *ɨ-i* and *ɨ-u* HRs in other contexts as well: an example is *hide-ude* 'ugly-pl; wet-pl'. Finally, *u* and *i* practically never rhyme; when they do (as in *pune-lumine* 'puts-lights'), it's in the context of an adjacent nasal. Although these frequency comparisons are informal and certain effects remain unexplored, they recapitulate many of the similarity relations in (8).

7. j-Ø mismatches

Romanian realizes word-final unstressed *i* as [j] in most contexts, after a vowel or a consonant. All poets studied allow C-Cj HRs like *nor-zborj* 'cloud; you-fly'. These are very frequent, as indicated by the figures in table (4): they represent more than a third of all HRs in the Shelley translation and about a quarter of the HRs in the translations of Briusov and Derjavin. In contrast, V-Vj HRs like *fade-baladej* 'insipid.pl-of the balad' are very rare. I have verified in the Labis corpus that this frequency difference between V-Vj and C-Cj HRs is not attributable to the frequency of word-final Cj and Vj. Labis has two V-Vj HRs and thirteen C-Cj HRs: figure (7) shows how these HRs compare to the overall frequency of Vj and Cj sequences in Labis's rhymes.

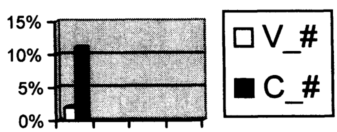


Figure 7: j-Ø HRs as a function of context in Labis
% = ratio of j-Ø HRs in a context to all j-final rhymes in that context

This data shows that Labis strongly prefers the C-Cj HRs to the V-Vj HRs: the relative frequency of line-final Cj and Vj is the same for Labis and thus cannot explain the rarity of his V-Vj HRs. The comparable difference between V-Vj and C-Cj HRs in all other corpora must be explained in the same way: poets tolerate a missing *j* much better if it is not adjacent to a vowel. Recall now the earlier comments on English *ez-ezd* HRs: Romanian C-Cj HRs are their direct counterparts[2]. They too are explained by the shared knowledge of the similarity ranking in (3).

8. Summary thus far

The last three sections have provided a sketch of the argument that establishes shared knowledge among Romanian poets for three similarity hierarchies – (3), (6), (8). It was also suggested that knowledge of these similarity relations is based on knowledge of perceptibility. We now consider how knowledge of the sound

[2] Romanian [j] is the surface realization of two very frequent suffixes: the plural and the 2nd pers. sing present. Other word final CC sequences are far less common in Romanian, hence the lesser opportunity for other C-Ø mismatches in word final position.

distributions in the Romanian lexicon interacts with the knowledge of perceptual similarity documented by the HR data.

9. Lexical knowledge does not affect similarity judgments

Recall that the major HR types (in (4)) do not correspond to alternations or distributional limitations. Below I list three productive Romanian processes: none gives rise to well-documented HRs. I discuss first the process highlighted .

(9) Productive segmental alternations lack HR counterparts

Process	Example	Restricted?	SR reflex? Attested?
t -> t^s / _ i	*frate / frat*^{s}*-ij* 'brother-brothers'	Derived environments	*lut-put*^{s} unattested *fete-bet*^{s}*e* unattested
s -> ʃ/ _i s -> ʃ/ _j	*pas/ paʃ -ij* 'step-the steps' *pas/ paʃ-j* 'step- steps'	Derived environments; native words Unrestricted	*kos - moʃ* unattested *waste-paʃte* very rare *deskiʃj-vis* very rare
ó ->wá/_C[-high] é ->eá/_C[-high]	*kot / kwat-e* 'elbow-elbows' *tem/ teamʌ* 'I fear-fear'	Derived environments; native words	*swarbʌ-tʃorbʌ* unattested *seamʌ-temʌ* unattested

I consider the hypothesis that the perceived similarity of a pair of strings is a function of its distributional status in the lexicon: in particular, I consider the possibility that members of an alternating pair will be perceived as functionally equivalent and that this will enhance their similarity (Hume and Johnson 2002) compared to non-alternating items. With this in mind, consider the extreme rarity of a HR type like *deskiʃj-vis* 'open.pl-dream', found only once in the entire corpus. The relevant facts are as follows: the palatalization of *sj* to [ʃj] is obligatory and thus surface [sj] is impossible. Further, most surface [ʃj] sequences – including the one in *deskiʃj* - correspond to /s-j/ inputs. Now when we unpack the effect of obligatory palatalization in *deskiʃj-vis,* we obtain a very popular HR type: the Cj-C variety discussed above. The question then is why the *ʃj –s* HRs like *deskiʃj-vis* are so rare, when HRs like *ʃj-ʃ* (like *uʃj–kʌuʃ* 'doors-muzzle') are among the most frequent. Perhaps the answer is to be found in the fact that any *ʃj-s* HR suffers from two mismatches: anteriority (*s* vs. *ʃ*) plus the added j-Ø difference. However HRs that cumulate two independently acceptable mismatches are very common in general. Just in Labis's collection – the poet who rhymes *deskiʃj* with *vis* - we find the following (with rhyming domains underlined): *strivind-simt, kind-strimt, atʃelor-dʒerurj, huruind-vint, bate-tatij, inimj-nimenj.* Thus the cumulative effect of feature differences does not explain why it is particularly *ʃj-s* that can't rhyme. The only explanation for this fact is the dissimilarity between *s-ʃ*: this pair almost never rhymes, in any context, *and their alternating status before i and j does not alter this fact.* We can conclude that alternation do not contribute, in this case, to the perception of similarity. The same can be shown for potential HRs generated by the automatic and robustly attested processes of coronal and velar palatalization before *j*: pairs like *t-tsj, z-ʒj, k-tʃj, g-dʒj* alternate but don't rhyme.

A distinct argument for the same point emerges when we consider the status of *i-ɨ* and *ɨ-u* in Romanian. The pair *i-ɨ* stands in near-complementary distribution: [ɨ] is rare, except after [r]; [i] is very frequent, except after [r], where it is rare in native forms. The pair *i-ɨ* alternate: in derived environments /i/ becomes [ɨ] after postvocalic [r]. The theme vowel of the 4th conjugation surfaces as [i] in *jub-i* 'love' but as [ɨ] in *ur-ɨ* 'hate'. Despite the distributional facts, HRs involving oral *i-ɨ* are rare, especially in 20th century poetry, as noted above. Compare now the pair *i-ɨ* with *ɨ-u*. Acoustically, the two pairs appear to be equally close (cf. 7), with the possible difference that *ɨ-u* is not differentiated by F3 values and thus might be perceived as more similar. But distributionally, there is a significant difference: unlike *i* and *ɨ*, *ɨ* and *u* neither alternate nor stand in near-complementarity. The question then is whether perceived similarity reflects the functional equivalence suggested by the distributional facts: if so, the alternating *i-ɨ* should rhyme more frequently. In the translation corpus we find 32 HRs involving the high pairs *i-ɨ* and *ɨ-u*: 28 of these occur before nasals (the context of nasalization) and 24 occur before CC (the context of vowel shortening). This is exactly as predicted by (8). Only 4 HRs involve oral vowels (*farfuzʌ-gizʌ, kʌtsuje-tʌmɨje, pis-rɨs, azvɨrle-turle*) and only one (*pis-rɨs*) involves an alternating vowel pair, *i-ɨ*. Thus functional equivalence does not affect similarity. If we consider all HRs involving a high vowel, nasalized or not, we find that one or both vowels are adjacent to [r] in 9/23 *u-ɨ* HRs and in 5/9 *i-ɨ* HRs. The conclusion is this: a neighboring [r] probably alters the similarity evaluation – perhaps because it lowers F2-F3 values - but does so equally for *u-ɨ* and *i-ɨ*. Recall though that [r] conditions only the distribution of the *i-ɨ* pair. Had distributional or alternating status affected the evaluation of similarity (as suggested by Hume and Johnson (2002)) we should have been able to observe this effect. We did not, as in the previous case.

10. Phonetic knowledge

In this paper nothing has been said about the source of similarity knowledge. The suggestion is that similarity judgments are rooted in the speaker/hearer's understanding that the perceptibility of contrasts varies in context. We can identify three perceptual factors that model the judgments discussed thus far: I'll refer to these as relative distance, the subset effect and cue duration. Relative distance refers to the case in which the contrasts x-y and z-w are realized on a shared auditory dimension D. If the distance in D between x-y is smaller than that between z-w then, all else equal, x-y is more similar. This characterization applies to the differences listed in (10), all of which correlate with the similarity judgments underlying HR use as discussed above.

(10) Relative distance effects: F2 distance
Δ_{F2}(u-i) > Δ_{F2}(u-ɨ), Δ_{F2}(i-ɨ);
Δ_{F2}(u-ɨ) > Δ_{F2} (ũ-ɨ̃); Δ_{F2}(i-ɨ) > Δ_{F2} (ĩ-ɨ̃)

The subset effect is a subcase of relative distance: if the pair x-y is differentiated by a proper subset of the acoustic dimensions differentiating z-w, then, all else equal, x-y is more similar. This characterization applies to the possible difference between *u-ɨ* and *i-ɨ*, whether oral or nasalized: the *u-ɨ* HRs are more frequent. M.R.Paraschivescu lacks *i-ɨ* HRs altogether. This may correlate with the absence of any F3 difference between *u-ɨ*.

Finally cue duration refers to case in which the pairs x-y and z-w are differentiated by some shared cue, which is however expressed over a longer period of time in z-w than in x-y. Then, all else equal, x-y is again the more similar. This characterization applies to the high vowel HRs (more frequent for shorter vowels) and to the post-nasal voicing HRs (the shorter postnasal voiceless closure yields a voiceless stop that is more similar to a voiced one).

I anticipate that an explicit model of perceptual similarity can be formulated by assigning weights to these factors and other comparable ones. This however must be left for future work.

11. Narrow lexical override

Suppose that naïve speakers of a language are aware of a relative similarity relation Δ(x-y)> Δ(z-w) and, as suggested earlier, they infer from this knowledge a correspondence ranking Corr(x-y)>>Corr(z-w). But suppose also that the language provides overt data showing that the opposite ranking is necessary to describe existing alternations. I have suggested thus far that knowledge of similarity exists, that it is shared among speakers even in the absence of lexical evidence for it, and that it is employed in loan adaptation. Now we must consider the cases in which the lexical evidence is at odds with the grammatical rankings inferred from knowledge of similarity. The case presented below is not entirely realistic but serves to illustrate the nature of the problem.

In Steriade and Zhang (2001), we argued that rhyming identity is computed by a set of correspondence conditions, which map elements of one rhyming domain (RD) onto corresponding elements of another line's domain. We further argued that the relative frequency of HR types reflects the rankings of different RD correspondence constraints: the frequent HRs violate low ranked RD constraints. The ranking itself is derived form similarity rankings, in accordance with the P-map hypothesis. Consider now the consequences of this proposal.

A common Romanian HR type involves post-tonic height mismatches, as in *rekínilor-kabínelor* 'of the sharks-of the cabins'. This HR, which looks like the rhyming equivalent of post-tonic reduction, violates Ident [±high]/$\sigma_{[+stress]}$_(RD). If HR frequency is an indication of the low rank of the relevant RD constraint then Ident[±high]/$\sigma_{[+stress]}$_ (RD) is low ranked. Consider now a conceivable but virtually unattested HR type, discussed earlier: *tot-hotsj* 'all-thieves'. This is the rhyming equivalent of assibilation and it violates Ident[±strident](RD). The very substantial frequency difference between "posttonic reduction" HRs and "assibilation" HRs suggests Ident[±strident](RD) >> Ident[±high]/$\sigma_{[+stress]}$_(RD).

The lexicon of Romanian shows that the opposite ranking holds among IO correspondence constraints: /t/ becomes [ts] before [j] as shown by alternations like *tot* 'all.sg' vs. plural *totsj* /tot-i/. The ranking is **tj* (a preliminary characterization of the phonotactic that triggers assibilation), **[+high]#* (the constraint causing final *i, u* to become glides), Ident[high]/$\sigma_{[+stress]}$_(IO) >>Ident[strident](IO). We focus on the ranking of correspondence needed to block a failed candidate like **tóte* for UR */toti*. Lowering posttonic /i/ in **tóte* will satisfy the phonotactic **tj* and will avoid the correspondence violations entailed by glide formation (*toti → totj*) and affrication (*totj → totsj*). To select *totsj* as the winning candidate we must then rely on Ident[high]/$\sigma_{[+stress]}$_(IO) >>Ident[strident](IO). But we have seen above that the opposite ranking holds for RD correspondence.

The question is this: if one of these correspondence rankings is inferred from knowledge of similarity, then clearly the other one is not. What then is the source of the other ranking? How do the two come to coexist in one grammar?

It is useful now to return to a point made at the beginning of this study about the nature of evidence for phonological relations. Learners of Romanian find out about the *ti → tsj* assibilation process by learning the words of Romanian. The same learners – if they are poets or readers of poetry – lack any comparably overt evidence about tolerable HRs. In that case then we must assume a hidden source of evidence: this, we proposed, is the knowledge of similarity. Under this interpretation the assibilation example shows a familiar fact: languages can acquire alternations – perhaps through the successive, telescoping action of local changes – whose synchronic result is to relate sounds viewed by speakers as highly dissimilar. It appears that, whatever role the judgments of similarity play in phonology, they do not inhibit such historical evolution: the dissimilarity of *t* and *ts* did not block the chain of events leading to assibilation in Romanian. This much is familiar.

There are also two less familiar points lurking here. First, if the learner infers rankings of correspondence from similarity relations, then it must also be that overt lexical evidence for the opposite ranking takes priority: if the P-map suggests Ident[±strident](IO)>>Ident[±high]/$\sigma_{[+stress]}$_(IO) and the lexical evidence shows the opposite, as we have seen, then the P-map based inference is blocked. The P-map then should be viewed as one inference mechanism, potentially overriden by other inference mechanisms, rather than as an absolute ban on certain linguistic relations. In this case, the lexical evidence of *t-ts* alternations overrides the similarity based inference that the two sounds cannot relate as correspondents.

However the Romanian evidence also shows that the lexical evidence for a given correspondence ranking is *narrowly interpreted* by learners: it is interpreted as bearing strictly on the lexical domain where the evidence originates. Romanian speakers do not view alternating pairs as more similar, just because they alternate: the simple fact that assibilation causes *t* to become *ts*, doesn't alter the similarity judgment, which HR selection reflects. The lexical evidence is narrowly interpreted: it requires a certain ranking of IO correspondence, but nothing further follows from this ranking. In contrast, it appears that similarity rankings are broadly construed, as bearing on all domains of correspondence – whether rhyming correspondence or the correspondence between a non-native input and its nativized correspondent - subject to lexical override.

References

Avram, A. 1963. Les Voyelles Roumaines [a] et [i] au point de vue acoustique. *Etudes et Recherches Linguistiques*, 14, 2, 165-179

Avram, A. 1968. Le Raport entre les Voyelles Neutres et la Nasalité. *Revue Roumaine de Linguistique*, 6, 567-573

Bell-Berti, F. 1993. Understanding Velic Motor Control. in M.Huffman and R.Krakow (eds.) *Nasals, Nasalization and the Velum*, Academic Press.

Eekman, T. 1974. *The Realm of Rhyme: A Study of Rime in the Poetry of the Slavs*, Verlag Adolf M.Hakkert, Amsterdam

Frisch, Stefan, Michael Broe and Janet Pierrehumbert 1997 Similarity and Phonotactis, *Rutgers Optimality Archive*.

Golston, Chris. 1994.The Phonology of Rhyme. ms. CSU Fresno

Halle, Morris. 1978. Knowledge unlearned and untaught . in M.Halle, J.Bresnan and G.Miller (eds.) *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press

Hayes, Bruce and Tanya Stivers. 1996 .A Phonetic Account of Post-Nasal Voicing. Ms. UCLA

Holtman, A. 1996. *A Generative Theoy of the Rhyme: An optimality Approach.* OTS, Utrecht

Hume, Elizabeth and Keith Johnson. 2002. The impact of phonological contrast on speech perception. To appear in *Proceedings of ICPhS.*

Newman, Paul. 2000. *The Hausa Language*, Yale University Press.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9.2:289-328.

Steriade, Donca 2001 "Directional asymmetries in place assimilation," in E.Hume and K.Johnson (eds.) *The Role of Speech Perception in Phonology*. New York: Academic Press. 2001.

Steriade, Donca. 2003. The phonology of perceptibility effects: the P-map and its consequences. in S.Inkelas and K.Hanson (eds.) *The Nature of the Word: in*

Honor of Paul Kiparsky. Cambridge, MA: MIT Press.

Steriade, Donca and Jie Zhang. 2001.Context dependent similarity. paper read at the 33rd meeting of the Chicago Linguistics Society.

Suteu, Valeriu .1963. Recherches sur la structure acoustique des voyelles roumaines .*Etudes et Recherches Linguistiques*.14, 2, 179-199

Wright, James. 1986.The Behavior of Nasalized Vowels in the Perceptual Vowel Space. in J.Ohala and J.Jaeger (eds.) *Experimental Phonology*. Academic Press.

Zwicky, Arnold 1976. Well, this rock and roll has got to stop. Junior's head is hard as a rock. *Chicago Linguistic Society* 12

Appendix: The translation corpus:

	Poet	Translator	Year	Total rhymes	HRs	HR frequency
B	Valeri Briusov	V. Kernbach	1961	761	131	17%
D	Gavril Derjavin	V.Teodorescu	1964	1386	156	11%
K	I.A.Krylov	T.Arghezi	1957	2258	15	7%
M	Vladimir Maiakovski	C.Teodorescu	1970	3230	315	10%
P	Aleksandr Pushkin	G.Lesnea et al.	1957	1865	46	.02%
S	P.B.Shelley	P.Solomon	1957	350	30	9%

The Non-neutralizing Nature of Hungarian Voicing Assimilation

ZOË TOFT* and WOUTER JANSEN**
*SOAS, University of London, **University of Groningen

0. Introduction[1]

In this paper we report on a pilot experiment designed to assess whether the process of regressive voicing assimilation in Hungarian applies in a categorically neutralizing manner, as implied by recent phonological analyses, e.g. Szigetvári (1998), Ritter (2000), Siptár & Törkenczy (2000), or whether it is better modelled as a gradient, phonetic rule (cf. Ernestus 2000 and Jansen 2001 on Dutch).

Our results, based on acoustic data from two speakers, indicate that Hungarian regressive voicing assimilation is not a neutralisation phenomenon. Whilst there is clear evidence of assimilation on the phonetic voicing of target obstruents and the preceding vowels, underlying distinctions in obstruents targeted by the process are still detectable in some of these features. In addition, we observe mismatches between the behaviour of voicing and segmental duration in obstruent clusters that contradict the predictions of the phonological accounts mentioned above.

Hungarian is a Uralic (Finno-Ugric, Ugric) language spoken by around 15 million people in Hungary and (as a minority language) in several of the surrounding states. The obstruent system of Hungarian is bifurcated into a set of *tense* and *lax* sounds in a manner that is similar to that of the surrounding Slavonic languages and Romanian (Kenesei *et al.* 1998; Siptár & Törkenczy 2000).[2] Unlike many of these languages, however, Hungarian maintains laryngeal contrast word finally before sonorants and prepausally, as in (1).

[1] This work was partially funded by award SSS/10484396 from the Students Award Agency for Scotland held by Zoë Toft, award 200-50-068 from Netherlands' Organisation for Scientific Research (NWO) held by Wouter Jansen and a stipend from the University of Tübingen granted to Wouter Jansen. First of all our thanks go to our test subjects. For their comments on (earlier versions of) this work we also thank John Nerbonne, Dicky Gilbers, members of the London Phonology Reading group, Péter Siptár, Mária Gósy and the audience at BLS 29. Special thanks go to Peter Sherwood for Hungarian inspiration. Of course, all errors remain our responsibility.

[2] The terms *tense/fortis*, *lax/lenis*, and [+/-tense] are used here as descriptive labels instead of *(phonologically) voiceless* vs. *(phonologically) voiced* and [+/-voice]. The latter represent more common terminology, but they obscure the difference between phonetic voicing (the acoustic result of vocal fold vibration) and lexical 'voicing' contrast, which is normally supported by a host of phonetic cues in addition to voicing.

In mixed tense + lax and lax + tense sequences, Hungarian word-final obstruents are subject to a process of *regressive voicing assimilation* (RVA). This process devoices lax obstruents followed by a tense plosive or fricative (2), and voices fortis obstruents before a lax plosive or fricative (3). As long as no pause intervenes, regressive assimilation is said to apply in sandhi clusters created by the morphology, by compounding (e.g. *rabszolga* 'slave', from *rab* 'prisoner', and *szolga* 'servant', in which underlying /bs/ assimilates to [ps]), and between independent words (e.g. *nyolc gyerek* 'eight children', with medial [d͡ʒɟ] from underlying /t͡sɟ/). The majority of researchers claim that this process is both obligatory and neutralising, and is not dependent on speech rate.[3]

(1)	nád	/na:d/	[na:d] *[na:t]	reed
	rág	/ra:g/	[ra:g]*[ra:k]	he chews
	láz	/la:z/	[la:z]*[la:s]	temperature
	lágy	/la:ɟ/	[la:ɟ]*[la:c]	soft
(2)	rabtól	/rɔb/+/to:l/	[rɔpto:l]	from (a) prisoner
	kádtól	/ka:d/+/to:l/	[ka:t:o:l]	from (a) bathtub
	habszifon	/hɔb/+/sifon/	[hɔpsifon]	cream-maker
	hadsereg	/hɔd/+/ʃɛrɛg/	[hɔtʃɛrɛg]	army
(3)	kalapban	/kɔlɔp/+/bɔn/	[kɔlɔb:ɔn]	in (a) hat
	kútban	/ku:t/+/bɔn/	[ku:dbɔn]	in (a) well
	szép zenész	/se:p/+/zɛne:s/	[se:bzɛne:s]	beautiful musician
	vak zenész	/vɔk/+/zɛne:s/	[vɔgzɛne:s]	blind musician

It would thus appear that the basic facts of laryngeal contrast and voicing assimilation in Hungarian are straightforward. Indeed Sauvageot (1951:27) typifies many commentators on this phenomenon when he states of Hungarian that "L'assimilation désonorisatrice et sonorisatrice est d'un mécanisme fort simple". The experiment reported below was designed to put this assertion to the test.

2. Methods

2.1 Speakers

Two subjects, K9 and M15, both female, took part in the experiment. The first speaker was 26 years old and had been a resident of Budapest for approximately 8 years at the time of recording. This subject grew up in Heves county and describes her own accent as 'standard Hungarian'. She is fluent in English. The second subject was 30 years old at the time of recording and although having

[3] See, for example, Hall (1944), Sauvageot (1951), Kálmán (1972), Lotz (1988), Kenesei, Vago and Fenvyesi (1988), Olsson (1992), and Siptár & Törkenczy (2000). A small number of authors, however, have suggested that voicing assimilation is not entirely obligatory: Kolmár (1821) and Vago (1980) suggest it is speech rate dependent, whilst Tompa (1961) believes that it can be suspended in foreign words and when the trigger consonant belongs to a contrastively stressed word.

moved frequently around Hungary describes her own accent as 'standard Hungarian'. She is fluent in English and French. Both speakers report normal hearing.

2.2 Stimuli

The stimuli for the experiment consisted of consonant clusters combining a /k, g, ʃ, ʒ/ C_1 and a /t, d, s, z/ or liquid /l/ or /r/ C_2. Stimuli containing a sonorant consonant were included to create baseline conditions for the comparison of the relative effects of fortis vs. lenis C_2 on the properties of a preceding obstruent. C_1 consonants were preceded by a long vowel or short vowel + glide sequence (phonetic diphthong) from the set /eː, aː, uː, ɔj/, or one of the following short vowels: /i, ɔ, o/. For reasons of space we will only consider clusters starting with /k, g/ here. The behaviour of clusters starting with a fricative is documented in Jansen & Toft (2002) and Jansen (submitted).

The clusters were located at subject noun + verb boundaries in carrier sentences. Subject + noun boundaries were chosen over other possible word boundary environments on grounds of the available carriers for C_1, which had to be similar in overall phonological make-up whilst exhibiting a robust contrast between /k, g/ (and therefore had to be unsuffixed). It was impossible to construct all carrier sentences according to the neutral word order for the propositions they expressed. This raised the possibility that the subjects would assign different prosodic structures to different stimulus sentences. However, all responses were pronounced with a F_0 peak on the subject noun carrying C_1, whilst the verb acting as the C_2 carrier never received any pitch prominence. Two sample stimuli appear in (4).

(4) a. A va*k* *d*arabolta a húst
/ɔ vɔ*k* *d*ɔrɔboltɔ ɔ huːʃt/
The blind man minced the meat

b. A rák zabálta a kis halacskát
/ɔ raː*k* *z*ɔbaːltɔ ɔ kis halat͡ʃkaːt/
The crab wolfed down the little fish

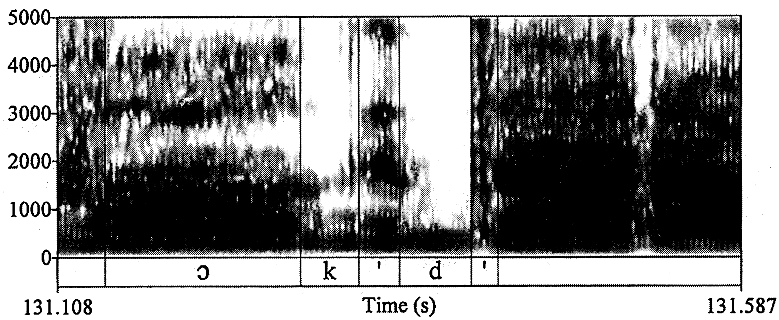
2.3 Procedure

The stimuli were presented to the subjects in a quasi-randomised order to avoid consecutive stimuli with identical consonant clusters. Each subject read the list of stimuli three times and was asked to read a stimulus again if she produced a mistake or hesitation that was clearly audible to the experimenter. In total, 2 (C_1) * 5 (C_2) * 6 (stimuli) * 3 (repetitions) * 2 (speakers) = 360 utterances with a plosive C_1 were recorded.

Recordings were made onto minidisk in a sound-proofed room using a Brüel and Kjær condenser microphone (Type 4165) and measuring amplifier (Type 2609), and digitised at 22.5 kHz. Segmentation and acoustic measurements were carried out using PRAAT version 4.0, as, for example, in (5). 36 utterances had to be discarded because they contained small speech errors, (hesitation) pauses between C_1 and C_2, or because the target cluster could not be internally segmented (plosive + plosive clusters with unreleased C_1). Note that our aim is to investigate the phonetic properties of Hungarian RVA, not the contexts in which it occurs. As most descriptions of the process assert that it is blocked by a physical pause this meant that it was important to exclude tokens with a pause intervening between C_1 and C_2: a failure to do so would have introduced a potential bias towards incomplete neutralisation in our corpus.

In addition, all 29 responses to stimuli that contained the C_1 carrier word *jog* 'law, right' were excluded from the analysis below because the mean duration of its vowel was felt to be exceptionally low. Tokens of this word were segmented such that nearly all of the F_2 fall from the palatal approximant into the mid back vowel was included with the latter, but nevertheless the mean duration of its vowel across C_2 environments was 75 ms, which is more than 2 standard deviations below the overall mean for the remaining lexically short vowels (cf. 8 below). This left 295 utterances for further analysis.

(5) Sample broad brand spectrogram of a /kd/ cluster. Speaker: K9 (female)



2.4 Segmentation and measurements

Segment boundaries and voicing intervals were determined by visual inspection of waveforms and broadband spectrograms. Voicing intervals were determined on the basis of periodicity in the waveform and the presence of a voice bar in the spectrogram. The closure and release phases of plosives were labelled separately (cf. 5). In the (few) instances in which the release phase of a plosive C_1 was visually completely obscured by the onset of a following fricative C_2, all of the aperiodic noise signal was assigned to the fricative, even if a release was (faintly) audible in it. The most important segmental boundaries were defined as follows:

- V_1 - plosive C_1: rapid decrease of higher frequency energy in the spectrum
- Plosive C_1/C_2 closure phase - plosive C_1/C_2 release phase: onset of release burst (defined as initial transient + following friction noise)
- Plosive C_1/C_2 release phase - C_2/V_2: end of release burst

- Onset and offset of C_2 fricative: onset and offset of aperiodic noise that could not be assigned to the release burst of a preceding stop

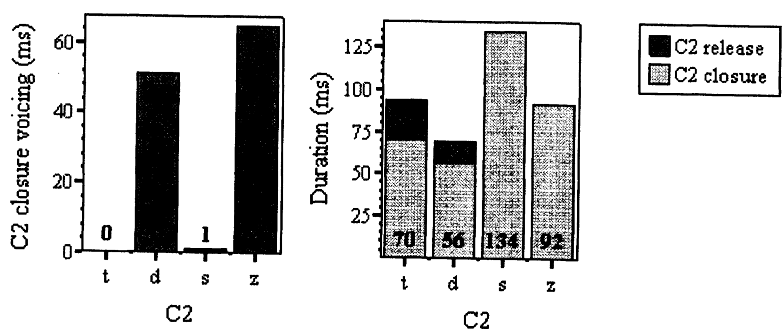
The following measurements were made on the basis of the hand-segmented speech samples: V_1 duration, C_1 closure duration, C_1 release duration and C_2 duration (closure and release separately for /t, d/) ; duration of voiced intervals during C_1 and C_2.[4] Obstruent duration, obstruent voicing duration and preceding vowel duration are all uncontroversial phonetic correlates of the tense-lax contrast in postvocalic contexts, although they are not always used to the same extent in different languages (see Keating 1984; Kingston & Diehl 1994 for overviews and references).

3. Results

3.1 Phonetic features of C_2 (the triggers)

Hungarian /t, d, s, z/ behave as might be expected on the basis of the phonetic literature on the realisation of tense-lax contrasts: the lax obstruents /t, d/ contrast with their tense counterparts in terms of voicing/VOT, and (closure and release) duration. The set of C_2 obstruents thus exhibits the same inverse behaviour of segmental duration and voicing duration that was found for C_1 obstruents in the baseline context (see below). In (6), C_2 voicing is expressed as closure voicing to facilitate comparison of fricatives (for which VOT is rarely used as a descriptive measure) and plosives.

(6) Duration and voicing of /t, d, s, z/. Left: closure voicing (ms); bottom: closure duration and (for plosives only) release duration (ms)



However, since it is a very common measure of plosive voicing we also calculated the VOTs for /t/ (23 ms, standard deviation = 7 ms) and /d/ (-51 ms, standard deviation = 11 ms).[5] These values put Hungarian firmly in the class of

[4] In addition, we measured F_0 at 10 ms intervals between 50 and 10 ms before C_1 onset and between 10 and 50 ms after C_2 (release) offset. For reasons of space we cannot comment on the behaviour of this parameter here, but note that it does not affect our overall conclusions.

[5] The VOT of /d/ may seem somewhat small in the light of the values reported elsewhere for lenis stops of the prevoiced type (e.g. Lisker & Abramson 1964; Keating 1984) but note that these

languages that realise the contrast between tense and lax plosives as *voiceless unaspirated* vs. *prevoiced* (cf. Lisker & Abramson 1964; Keating 1984), and this is consistent with the literature on Hungarian.

T-tests confirm that all observed differences are highly significant: plosive C_2 closure duration, $t(115) = 5.81$, $p <. 001$; plosive C_2 release duration, $t(115) = 7.81$, $p < .001$; plosive C_2 VOT, $t(115) = 43.45$, $p <.001$; fricative C_2 duration, $t(114) = 14.71$, $p < .001$; fricative C_2 voicing, $t(114) = -17.94$, $p < .001$.

3.2 Duration and voicing of C_1 (the targets)

The table in (7) below provides the overall duration and the duration of the voiced intervals for the closure and release phases of /k, g/ across C_2 contexts. The patterning of C_1 phonetic voicing shows unmistakable signs of regressive voicing assimilation to both tense and lax obstruents, and there is evidence that C_1 release duration is subject to assimilation too, albeit to a lesser extent. However, it is hard to interpret the behaviour of C_1 closure duration in assimilatory terms, and there is some indication that both this parameter and C_1 (release) voicing maintain the underlying distinction between /k/ and /g/ when followed by a lax obstruent.

Consider first the phonetic contrast between the tense and lax velar stop in the baseline pre-liquid context. Here, all of the parameters represented in (7) contribute to the phonetic expression of the lexical opposition between /k/ and /g/ in the expected fashion: the closure and release of /k/ are considerably longer than those of /g/, whilst its voiced intervals are shorter (the difference in overall voicing is 28 ms). T-tests confirm that the differences between the means for /kL/ and /gL/ are statistically significant: C_1 closure duration, $t(60) = 11.16$, $p < .001$; C_1 release duration, $t(60) = 4.62$, $p < .001$; C_1 closure voicing, $t(60) = -4.83$, $p < .001$; C_1 release voicing, $t(60) = -8.16$, $p < .001$; C_1 overall voicing, $t(60) = -11.52$, $p < .001$.

Of these parameters, those related to C_1 voicing show the clearest sign of assimilation to a following obstruent. Thus, before the lax obstruents /d, z/ the voicing of /k/ increases vis-à-vis its voicing in the baseline context, and there is a decrease in the voicing of /g/ compared with the relevant baseline values. As far as obstruent clusters are concerned, there appears to be little trace of the underlying distinction between /k/ and /g/: differences within the minimal pairs /kt/ - /gt/, /kd/-/gd/, /ks/-/gs/ are all equal to or smaller than 4 ms, whilst /g/ has 6 ms more closure voicing than /k/ when /z/ follows. Before a tense C_2 differences in C_1 release voicing are also virtually neutralised, but when a lax obstruent follows /k/ has 8 ms less voicing than /g/. Added to the differences in voicing during the closure stage this yields differences in overall voicing between the tense and lax velar stop of 10 ms before /d/, and 14 ms when the lax fricative /z/ follows. This suggests that whilst Hungarian word final stops are subject to RVA, this process is not completely neutralising across C_2 environments.

---

values often concern postpausal or utterance initial plosives. On aerodynamic grounds the presence of a preceding obstruent is likely to have had some negative influence on the amount of voicing of /d, z/, whilst the relatively short mean closure duration of /d/ naturally places an upper bound on its mean VOT.

A series of ANOVAs was carried out on the C_1 voicing data to gauge to what extent these impressionistic observations are supported by statistical tests. First, a three-way ANOVA for *C_1 laryngeal specification* (/k/ vs. /g/) * *C_2 laryngeal specification* (/t, s/ vs. /d, z/) * *C_2 manner* (/t, d/ vs. /d, z/) was performed on the C_1 closure voicing data (pre-liquid baseline environments excluded). This ANOVA reveals a highly significant effect of *C_2 laryngeal specification*, $F(1,225) = 69.16$, $p < .001$, but only a weakly significant effect of *C_1 laryngeal specification*, $F(1,225) = 5.74$, $p < .02$, which implies that regressive assimilation is near-neutralising with regard to this parameter (there were no other significant effects).

(7) Duration and voicing of C_1 (all measures in ms; standard deviations in brackets)

C_1C_2	C_1 closure duration (ms)	C_1 release duration (ms)	C_1 closure voicing (ms)	C_1 release voicing (ms)	N
/kt/	57 (13)	31 (10)	28 (7)	0 (2)	31
/kd/	57 (19)	28 (8)	40 (11)	17 (11)	31
/ks/	56 (10)	16 (7)	27 (7)	0 (0)	30
/kz/	59 (20)	22 (14)	37 (11)	14 (15)	29
/kL/	75 (12)	31 (11)	29 (9)	2 (3)	32
/gt/	52 (12)	32 (10)	29 (11)	1 (5)	30
/gd/	43 (11)	24 (8)	43 (11)	25 (8)	25
/gs/	50 (11)	14 (10)	31 (14)	0 (0)	28
/gz/	44 (10)	23 (12)	43 (10)	22 (13)	29
/gL/	42 (11)	19 (10)	41 (12)	17 (10)	30

However, three-way ANOVAs for *C_1 laryngeal specification* * *C_2 laryngeal specification* * *C_2 manner* on the C_1 release voicing and C_1 overall voicing data (baseline environment excluded) show stronger effects of *C_1 laryngeal specification* which indicate that the underlying distinction between /k/ and /g/ is retained before lax obstruents even if RVA does apply. Thus, the ANOVA on the C_1 release voicing data shows significant main effects of *C_1 laryngeal specification*, $F(1,225) = 11.89$, $p < .005$, *C2 laryngeal specification*, $F(1,225) = 284.63$, $p < .001$, and an interaction of *C_1 laryngeal specification* * *C_2 laryngeal specification*, $F(1,225) = 10.04$, $p < .005$. Unsurprisingly, the ANOVA on the C_1 overall voicing data shows the same set of significant effects: *C_1 laryngeal specification*, $F(1,225) = 16.10$, $p <.001$; *C_2 laryngeal specification*, $F(1,225) = 289.63$, $p < .001$; C1 laryngeal specification * *C_2 laryngeal specification*, $F(1,225) = 6.38$, $p <.015$.

As before, the effects of C_2 laryngeal specification indicate that the voicing of Hungarian word-final stops is subject to assimilation to a following obstruent, but the highly significant effects of *C_1 laryngeal specification* indicate that assimilation does not erase all the distinctions between underlying /k/ and /g/. The latter conclusion must be qualified somewhat in light of the significant interactions of the two main factors, which are most likely due to the fact that there are no differences in C_1 release and overall voicing between underlyingly tense and lax velar stops before /t, s/ whilst there are differences when a lax obstruent follows. These interactions support the idea that Hungarian RVA is non-neutralising in a subset of environments (i.e. before lax obstruents only). Finally, the absence of any effect related to C_2 manner of articulation indicates that Hungarian RVA is symmetric with respect to the plosive-fricative distinction: both classes of obstruent appear to trigger assimilation in equal measure.[6]

Now if regressive assimilation affects C_1 (release and closure) duration as well as C_1 voicing, this would result in a decrease of the duration of /k/ before /d, z/ vis-à-vis the baseline environment, and an increase in the duration of /g/ before the tense obstruents /t, s/. As can be gleaned from (7), this expectation is only borne out by the C_1 closure duration data. First, the closure stage of /k/ is considerably shortened before both lax and tense obstruents relative to the 75 ms observed in the pre-liquid context, and second, the [±tense] value of an obstruent seems to have little impact on the closure duration of a preceding /k/. Only the behaviour of /g/, which exhibits lengthening before /t, s/ is suggestive of regressive assimilation. Another partial parallel between C_1 closure duration and C_1 (overall) voicing is that the underlying distinction between /k/ and /g/ is better preserved before /d, z/ where there is a 14-15 ms contrast, than before /t, s/, where the difference in closure duration is 4-5 ms. Nevertheless the behaviour of C_1 closure duration fails to match that of C_1 (closure) voicing, which shows tense-symmetric assimilation to a following obstruent. This indicates that these two acoustic parameters reflect distinct (phonological and/or articulatory) control mechanisms.

Next, the first generalisation to emerge from (7) concerning the behaviour of C_1 release duration is that the release of /k/ and /g/ is relatively short before fricatives. As hinted in section 2 above, this is likely to be a labelling artefact caused by the overlap of release and frication noise in the acoustic signal, and it therefore seems safer to exclude cases involving a fricative C_2 from further analysis. This leaves the obstruent sequences ending in a /d/ or /t/. The data for these clusters seem to show a (small) effect of regressive assimilation: on average the release of /k, g/ is 8 shorter before /d/ than before /t/.

A three-way ANOVA for *C_1 laryngeal specification* * *C_2 laryngeal specification* * *C_2 manner* on the C_1 closure duration data (baseline context excluded) confirms the impression that this parameter behaves differently from the C_1 voicing measures: it reveals a highly significant main effect of *C_1 laryngeal specification*, $F(1,225) = 31.79$, $p < .001$ as well as an interaction between *C_1 laryngeal specification* and *C_2 laryngeal specification*, $F(1,225) = 6.84$, $p < .001$, but no other significant effects. The main effect of *C_1 laryngeal specification* indicates that C_1 closure duration maintains the distinction between /k/ and /g/, whilst the absence of a main effect of *C_2 laryngeal specification* supports the observation that, unlike C_1 voicing, this phonetic feature is not subject to RVA. Finally, the interaction between *C_1 laryngeal specification* and *C_2 laryngeal specification* is likely to be due to the fact that the difference

[6] Using (closure and release) *voicing ratio* (voicing duration divided by overall duration) instead of absolute voicing duration as an index of voicing assimilation leads to very similar conclusions.

between /k/ and /g/ is larger before tense obstruents than before lax obstruents, and this, therefore, confirms the impression that incomplete neutralisation of the contrast between these two plosives is again restricted to one particular class of C_2 obstruents.

By contrast, a two-way ANOVA for *C_1 laryngeal specification * C_2 laryngeal specification* on the C_1 release duration data (only cases with a C2 plosive included) yields a highly significant main effect of *C_2 laryngeal specification*, $F(1,113) = 11.87$, $p < .001$ but no other effects, indicating that this variable is subject to neutralising regressive voicing assimilation.

3.3 V_1 duration

The table in (8) presents mean values for V_1 duration by C_1 + C_2 context and separately for lexically long (/eː, aː, uː, ɔj/) and short (/i, ɔ, o/) vowels. Although long and short vowels were balanced across C_1 and C_2 contexts in the stimulus set, the fact that a number of responses had to be discarded resulted in some slight imbalances in the corpus. For example, there were more tokens of C_1 = /g/ + C_2 = /L/ preceded by a short vowel (12) than tokens of C_1 = /k/ + C_2 = /L/ (14). Since this might distort the value for V_1 duration in the second context to an artificially low value, they are presented separately here.

The behaviour of V_1 duration is different from the phonetic features discussed above in that at least in the corpus under consideration here, and with respect to plosive C_1 consonants, it is the only correlate of [tense] that assimilates in neutralising fashion across C_2 contexts.[7]

(8) V_1 duration (ms) for lexically short and long vowels/diphthongs (standard deviations in brackets)

C_1C_2	V_1 duration (ms)			
	Short vowels	N	Long vowels	N
/kt/	101 (14)	14	128 (20)	17
/kd/	115 (11)	16	138 (27)	15
/ks/	106 (14)	13	135 (22)	17
/kz/	109 (12)	12	134 (35)	17
/kL/	98(13)	14	125 (22)	18
/gt/	105 (20)	12	129 (21)	18
/gd/	115 (13)	8	148 (34)	17
/gs/	105 (15)	11	137 (23)	17
/gz/	114 (7)	11	148 (28)	17
/gL/	111 (8)	12	153 (33)	18

In the baseline pre-liquid environment, the speech of our two subjects conforms to the near-universal generalisation that vowels are longer before lax obstruents than before tense obstruents (Chen 1970; Kluender et al. 1988): note

[7] This conclusion may have to be modified once a larger set of contexts and speakers is taken into account. See Jansen & Toft 2002 and Jansen (submitted) for details

that the difference appears to be somewhat larger for lexically long vowels (28 ms) than for short vowels (13 ms). A two-way ANOVA for *V_1 lexical length* (short vowels vs. long vowels) * *C_1 laryngeal specification* confirms that lexical distinctions in vowel length and the distinction between /k/ and /g/ both affect the phonetic duration of V_1. There are highly significant main effects of *V_1 lexical length*, $F(1,58) = 34.69$, $p < .001$ and *C_1 laryngeal specification*, $F(1,58) = 12.85$, $p < .005$, but there is no significant interaction, which indicates that contrary to appearances, lexically long and short vowels mark the [tense] contrast on following obstruents in essentially the same way.

However, when C_1 is followed by an obstruent, V_1 duration seems to signal [tense]-contrasts in C_2 rather than in C_1: except in /k/ + fricative sequences V_1 is consistently longer before a lax C_2 than before a tense C_2. It is true that, on average, lexically long vowels are somewhat longer when followed by /g/ (140 ms) than when followed by /k/ (134 ms), but a four-way ANOVA for *V_1 lexical length* * *C_1 laryngeal specification* * *C_2 laryngeal specification* * *C_2 manner of articulation* shows that this difference is not statistically significant. The only significant effects revealed by this analysis are main effects of *V_1 lexical length*, $F(1,217) = 91.05$, $p < .001$, and *C_2 laryngeal specification*, $F(1,217) = 9.67$, $p < .005$, which indicates that before obstruent clusters, phonetic vowel duration is controlled by phonological length and regressive assimilation.

This result might be interpreted as evidence for the idea that there is a unitary mechanism of regressive voicing assimilation (or more appropriately tense-assimilation) that governs the behaviour of all phonetic exponents of [tense]. Current generative analyses of regressive voicing assimilation propose that this unitary mechanism consist of the spreading or 'agreement' of the phonological feature that represents the contrast between tense and lax obstruents at the lexical level (e.g. autosegmental [voice]). However, such analyses predict that all the individual cues to [tense] behave in parallel, and this prediction is not supported by the current data: C_1 (overall) voicing assimilates across contexts, but in non-neutralising fashion before lax obstruents whereas V_1 duration assimilates in neutralising fashion across contexts and C_1 closure duration does not assimilate at all.

4. Discussion and conclusions

Hungarian RVA has typically been described as a process that neutralises underlying fortis-lenis distinctions in the obstruents it targets. The aim of the pilot experiment reported in this paper was to assess to what degree this view can be supported by quantitative acoustic data. The results of this experiment indicate that voicing and the closure duration of plosives partially preserve underlying tense-lax distinctions in assimilation, despite the presence of a clear assimilation effect on voicing and on the duration of preceding vowels. Furthermore, different cues participate in different ways in the assimilation process. These observations raise questions about the viability of models that represent RVA as a single lexical-feature-spreading operation and support the more phonetic view of the phenomenon as alluded to in our introduction. Thus, this study contributes to a growing body of instrumental research demonstrating incomplete neutralisation effects of voicing assimilation rules (O. Thorsen 1966; N. Thorsen 1971; Charles-Luce 1993; Burton & Robblee 1997; Barry & Teifour 1999).

Nevertheless, we want to emphasise that the data reported above should be treated as preliminary observations that need to be supported by a larger-scale study of the phonetics of Hungarian obstruent clusters. Moreover, a number of empirical issues need to be explored before a more robust verdict can be delivered on the nature of the Hungarian regressive voicing assimilation rule. We think it is possible, for example, that this rule may be sensitive to juncture strength and operates in a neutralising fashion at relatively weak morphosyntactic or prosodic boundaries (e.g. in sandhi clusters created by the morphology). It is also important to establish whether the residual traces of underlying tense-lax distinctions in assimilation targets that we observe in the acoustic signal are perceived as cues to lexical phonological representations by listeners.

References

Barry, M. & R. Teifour. 1999. Temporal patterns in Arabic voicing assimilation. *Proceedings of the XIVth International Congress of Phonetic Sciences* (San Francisco). Vol 3: 2429-2432.

Burton, M. & K. Robblee. 1997. A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics* 25: 97-114.

Charles-Luce, J. 1993. The effects of semantic context on voicing neutralisation. *Phonetica* 50: 28-43.

Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22: 129-159.

Ernestus, M. 2000. *Voice assimilation and Segment Reduction in casual Dutch.* PhD dissertation, Free University of Amsterdam/Holland Institute of Generative Linguistics.

Hall, R. 1944. *Hungarian Grammar.* (Language Monograph 21). Baltimore: Linguistic Society of America.

Jansen, W. 2001. Phonetic voice and phonological assimilation in the Germanic languages. In *CLS 37: The Panels. Papers from the 37th Meeting of the Chicago Linguistic Society.* Vol. 2. Chicago: Chicago Linguistic Society.

Jansen, W. submitted. *The Phonetics and Phonology of Laryngeal Contrast in Germanic.* (Working title). PhD dissertation, University of Groningen.

Jansen, W. & Z. Toft 2002. On sounds that like to be paired (after all): an acoustic investigation of Hungarian voicing assimilation. *SOAS Working Papers in Linguistics* 12: 19-52.

Kálmán, B. 1972. Hungarian historical phonology. In L. Benkő & S. Imre (eds.) *The Hungarian Language.* The Hague: Mouton

Keating, P. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60: 286-319.

Kenesei, I., R. Vago & A. Fenyvesi. 1998. *Hungarian.* (Routledge Descriptive Grammar). London: Routledge.

Kingston, J. & R. Diehl. 1994. Phonetic knowledge. *Language* 70: 419-454.

Kluender, K., R. Diehl, & B. Wright. 1988. Vowel length differences before voiced and voiceless consonants: an auditory explanation. *Journal of Phonetics* 16: 153-169.

Kolmár, J. 1821. Próbatétel a Magyar Helyesirás Philosophiájára. In *Jutalom feleletek a' Magyar Nyelvről, a Magyar Nemzeti Muzeum 1815, 1816, 1817 esztendei kérdéseire*. Horvát István: Pest.

Lisker, L. & A. Abramson. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20: 384-422.

Lotz, J. 1988. *Das Ungarische Sprachsystem*. Bloomington, Indiana: Eurolingua.

Olsson, M. 1992. *Hungarian phonology and morphology*. Lund: Lund University Press.

Ritter, N. 2000. Hungarian voicing assimilation revisited in Head-driven Phonology. *Approaches to Hungarian* 7: 23-50.

Sauvageot, A. 1951. *Esquisse de la langue hongroise*. Paris: C. Klincksieck.

Siptár, P. & M. Törkenczy. 2000. *The Phonology of Hungarian*. Oxford: Clarendon.

Szigetvári, P. 1998. Voice assimilation in Hungarian: the hitches. In L. Varga (ed.) *The Even Yearbook 3. ELTE SEAS Working Papers in Linguistics*. Budapest: Department of English Linguistics, Eötvös Loránd University.

Thorsen, N. 1971. Voicing in British English t and d in contact with other consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen* 5:1-39.

Thorsen, O. 1966. Voice assimilation of stop consonants and fricatives in French and its relation to sound duration and intra-oral air pressure. *Annual Report of the Institute of Phonetics, University of Copenhagen* 1: 67-76.

Tompa, J. (ed.) 1961. *A mai magyar nyelv rendszere; Leíró nyelvtan*. Kötét 1. Budapest: Akadémiai kiadó.

Vago, R. 1980. *The Sound Pattern of Hungarian*. Washington, D.C.: Georgetown University Press.

Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street
Russell Square
London WC1H 0XG
United Kingdom

zt1@soas.ac.uk

Department of Linguistics
University of Groningen
P.O. Box 716
9700 AS Groningen
The Netherlands

w.jansen@let.rug.nl

Self-Organization and Categorical Behavior in Phonology

ANDREW WEDEL
University of California, Santa Cruz

0. Introduction

One of the most salient properties of phonological systems are their very consistency. This paper describes work suggesting that such categorical behavior in phonological systems can be modeled as an emergent property resulting from self-organization (Nicolas and Prigogine 1977) within an iterating, richly-specified lexicon.

1 Generative accounts of categorical behavior in phonology

Rule systems (e.g. Chomsky and Halle 1968) account for consistent patterns in surface form in given contexts through context-sensitive rewrite rules that act upon underlying forms stored in the mental lexicon. Categorical behavior follows from the stipulated mechanism of rule application: if the structural conditions of the relevant rule and/or constraint are met, a rule must apply, otherwise it must not. Further, if rules can possibly conflict, they are ordered with respect to one another, allowing one of the rules to consistently determine the outcome of all such conflicts.

Where rules systems rely on application of inviolable rules to derive input-output relations, in Optimality Theory (OT, Prince and Smolensky 1993) lexical inputs are mapped to optimal outputs through the satisfaction of violable, ranked constraints on output form. Because the ranking of constraints is fixed, any input strings that share a relevant set of morpho-phonological properties will exhibit an analogous mapping relationship to their optimal outputs. The ranking, or *dominance* relation between two constraints is made apparent when an optimal output violates one constraint in order to satisfy the other. In this paper, I will use the term 'dominance' beyond its strict OT usage to refer to a ranking relation between universal constraints, to refer to the more general categorical satisfaction of one surface pattern at the expense of another within a language.

In addition to simple dominance relations between conflicting patterns, grammars often exhibit a higher-order kind of dominance that becomes apparent when multiple patterns collide in one output form, in which the result of a conflict between multiple patterns tends to follow the results of the component pair-wise pattern conflicts. OT accommodates this observation through the stipulation that constraint dominance is *strict*, that is, that output candidates satisfying a higher ranked constraint will always win over candidates violating that constraint, while satisfying any number of lower ranked constraints (Prince and Smolensky 1993). In what follows, I will generalize the OT term 'strict dominance' beyond its strict OT usage to refer to the persistence of pattern dominance in the face of conflict with multiple, potentially cooperating patterns.

The aim of the work described here is to show that grammatical relations exhibiting both dominance and strict dominance emerge spontaneously within psycholinguistically plausible models of the lexicon (Tenpenny 1995, Pierrehumbert 2001). The work will be based in a general model of language production and processing that satisfies the following two general conditions:

1) The lexicon is able to simultaneously store detailed information derived from individual events, and multiple categorial abstractions over that information.
2) The mechanism for assembling production targets for a given linguistic element causes such targets to be biased toward the form of other, similar linguistic elements.

The first condition provides a mechanism for low-level phonetic effects to build up and drive category evolution over time, while the second creates a leveling tendency within the lexicon, which will be shown to promote evolution of lexicons exhibiting dominance and strict-dominance patterns in their grammatical relations.

2. Exemplar models of categorization and the lexicon

The assumption of a fundamental distinction between general, abstract knowledge and specific, episodic memory has a long tradition in the psychological literature on categorization. In recent decades however, research has repeatedly found that subjects retain access to highly detailed, episodic memories of an event for a surprisingly long time (reviewed in Johnson 1997), and make use of these memories when carrying out tasks thought to require only general knowledge (see Tenpenny 1995 and references therein). As a consequence, a class of new theories has developed which locate specific, episodic memories at the core of categorization processes (Hintzman 1986). While such theories do not deny that generalizations exist, they begin from the hypothesis that abstract knowledge has no special status relative to specific knowledge, and that abstract knowledge does not require a form of representation distinct from that encoding specific memories. In the last decade, exemplar theory has been extended to the domain of language by linguists and psycholinguists interested in categorization phenomena both in perception (Goldinger 1996, Johnson 1997) and in production (Pierrehumbert 2001).

In exemplar models, each category is defined by a 'cloud' of remembered tokens, or exemplars, that have been tagged as belonging to that category. Exemplars are organized within the category by similarity across any salient dimension, producing internal structure in category-space; a given exemplar may therefore contribute to many categories simultaneously (reviewed in Tenpenny 1995, Hintzman 1986, Goldinger 1996, 2000, Pierrehumbert 2001, 2002)

3. A General Framework

The general framework structuring this work begins with the assumption that every incoming chunk of information can potentially be broken apart and categorized at many levels, such that phoneme level categories coexist with, and are cross-referenced with, categories made up of smaller and larger sequences. In the model, an intention to produce a form, e.g. a word, results in spreading activation of not only previously categorized exemplars of that form itself, but all other exemplars that contain or are contained by it, weighted by strength and number of connections (cf. Pierrehumbert 2001). These activations all may contribute to a production target for that form.

The nature, timing and degree of interactions between categories is left intentionally vague in the present account; for our purposes, the only important feature of the model that derives from these interactions is that a production target be influenced by multiple, cross-categorized elements. This feature has the following important consequence: provided that categorized elements retain their individual values on the dimension that led to their categorization, as in exemplar theory, production of a new category element under any influence of consensus within pre-existing elements will result in 'blending inheritance', with consequent reversion to the mean of the category (Abler 1997, Pierrehumbert 2001).

The resulting tendency to assemble outputs that conform to a mean over previously stored forms constitutes, in effect, a form of analogical pressure, operating in this case at the level of phonological production targets. It is this analogical pressure that will be shown in later sections to interact with external biases to produce familiar phonological patterns over time.

4. Properties of analogical systems

A system can be described as analogical when the future behavior of a system element is biased towards the present behavior of other system elements. There are two properties of such systems that will concern us here:

1) Gradient patterns in behavior are unstable; persistent bias towards similarity between elements promotes sharp boundaries in behavior.
2) In the absence of forces that maintain or add difference in a finite system , all system elements will eventually come to exhibit identical behavior.

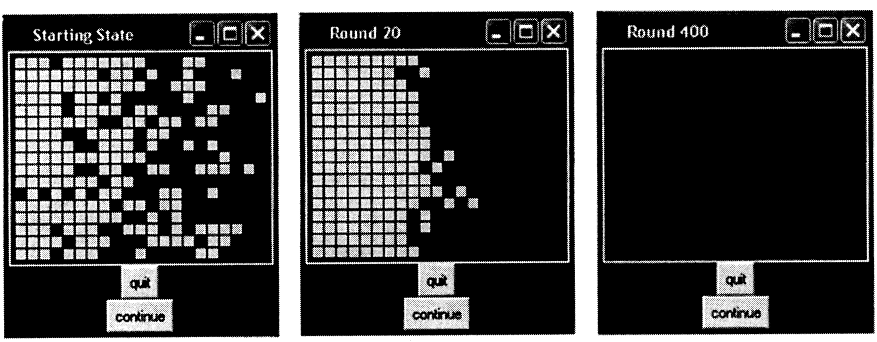
Both properties directly derive from the bias toward similarity. To illustrate this, I show stages in the evolution of a simple cellular automata program (Figure 1A-C). At the start of the simulation, each square in a field is randomly assigned a light or dark shade of gray, with the caveat that there is a gradiently higher probability of being assigned dark gray from left to right (Fig. 1A). In each subsequent round of the simulation, each square has a small probability of changing its shade. Crucially, however, the choice of shade is stochastically biased towards that of squares in the vicinity, with more weight given to nearer neighbors' shades. The result is that the initially gradient light-dark pattern rapidly becomes more categorical, as seen after 20 rounds in Figure 1B.

Note that this segregation occurs despite the fact that there is no explicit directive to do so in the conditions of the simulation: there is only a tendency for a square to become more like its neighbors. This creates a basin of attraction characterized by shade-identity, such that any changes tend to expand consensus neighborhoods at the expense of mixed neighborhoods. The greatest reduction in boundary between two neighborhoods in this 2-dimensional field is a straight line, which the simulation continues to approach beyond the point seen in Figure 2B.

Because the squares along the boundary have mixed neighbors, they continue to alternate, allowing the boundary to migrate back and forth across the field. When it approaches an edge, the minority shade begins to lose to the majority, and may disappear for good as occurred here on round 400 (Fig. 1C). The initial segregation of shades into two regions, followed by eventual loss of one shade, illustrates a fundamental property of analogical systems: difference is unstable. Here we see that shade difference is initially minimized locally by segregating shades together, but is eventually wiped out globally as well.

Figure 1

A. B. C.



5. Simulation of pattern formation within the framework

The general framework introduced in section 3 is an analogical system, in which production targets for a lexical entry are constructed under the influence of a large set of activated exemplars sharing sequences with that lexical category. The actual productions of lexical entries are consequently biased in the direction of a mean for the lexicon. Re-storage of such productions as exemplars in the lexicon results in a feedback loop, continually updating the lexicon with new exemplars that are more similar to one another.

However, no lexicon ever comes to eventually consist of one word – on the contrary, diachronic change gives the impression of a constantly shifting equilibrium. From the point of view of this framework, such an equilibrium can only be maintained through forces that support or introduce difference within the lexicon. The source of difference that will be relevant in this paper is located in context-sensitive biases in articulation and perception – precisely those gradient, phonetic-level tendencies that so many linguists have proposed form the raw material for grammaticalization (reviewed in Blevins to appear)[1]. The combination of these two general features produces a a system in which idiosyncratic behavior steadily seeps into the lexicon under the influence of low-level performance biases, while reversion to the mean steadily works to reduce that idiosyncrasy. We may predict then that in the ensuing shifting equilibrium, gradient patterns introduced by chance or through consistent biases in performance may occasionally be converted into categorical patterns in the lexicon. The following section introduces a simulation architecture designed to allow testing of this hypothesis within the model.

6. Architecture of the simulation

The architecture presented here is not meant in any way to be taken as a claim about the actual physical functioning of the mental lexicon and production mechanism, but represents rather a serial mechanism for reproducing three key properties that the lexicon and production system are hypothesized to possess, as described above. These properties are

- Multiple nested categorizations of events.
- Influence of multiple categories in assembly of production targets.
- Storage of phonetic detail in memory.

The simulation consists of three parts: a lexicon consisting of abstract lexical entries and corresponding exemplars, an implementation mechanism that uses the information inherent in the lexicon to produce production targets, and a performance filter that introduces biases in actual output form.

In the lexicon, lexical entries are split into two levels, the first an underlying form composed of ordered, abstract phonological categories, and the second a set of more phonetically detailed, remembered exemplars of previous outputs from that lexical entry[2].

---

[1] Pressure to maintain contrast for communicative efficiency provides a distinct set of forces working against leveling pressure in the lexicon. See Wedel (forthcoming) for further discussion.
[2] In the current simulation model, lexical entries are fixed, while exemplars are free to vary under the influence of performance biases and pressure from other exemplars. The disconnect between abstract lexical entry and exemplar cloud is necessary at this point because the lack of any force maintaining contrast results in a collapse toward uniformity if lexical entries are not fixed. See

To simulate a web of lexical entries interacting in the assembly of a given lexical entry's production target, all lexical entry substrings that show any contiguous featural match to substrings in the given entry are identified. The set of relations between those substrings and their reflexes in corresponding exemplars are then summed to produce a probability that a given phonological category in the lexical entry will be associated with a particular reflex in the production target. These summed relations represent a set of weighted implications inherent within the lexicon, each stating essentially, an input substring X should surface as the substring Y. The set of weighted implications is then used to construct a set of possible production targets, where the optimal target enjoys the greatest support over the entire weighted implication set. Once assembled, the production target is sent to Performance.

Performance contains a set biases comprising feature specifications identifying marked sequences, a possible change, and the likelihood of that change occurring. A bias against complex onsets, for example, would be activated whenever a production target contains the feature sequence [onset], [C], and would possibly exchange the feature [coda] for [onset]. The likelihood of this change occurring is set in this simulations here at 10%, meaning that if the production target constructed by the lexicon is [u.bli], there is a 10% chance that it will be articulated as [ub.li] instead. This output is stored in a temporary buffer, and the lexicon begins the process again for another entry. After every entry in the lexicon has produced all of its requisite outputs, the exemplars for each entry are overwritten by the corresponding outputs from the temporary buffer. The process then begins again, with new exemplars for each entry. For a more detailed description of the algorithm, see Wedel (forthcoming).

7. Analogical structure results in categorical behavior

The first simulation shown runs over a lexicon consisting of 10 entries of the form VCCV (e.g., /ubli/), seeded with an equivalent number of exemplars with syllable boundaries before and after the first consonant. This simulation includes no performance biases that filter outputs, such that production targets are produced without modification. The symmetry in the syllabification of the seed exemplars means that in the initial round of the simulation, the set of implications derived from the lexicon for each lexical entry will provide no advantage to one syllabification over the other, resulting in random syllabification choices. However, the moment there is a numerical advantage of one syllabification pattern over the other in some lexical entry, the system will tend to exaggerate that pattern, spreading it first to those lexical entries most similar to that originating the bias, and from there to the entire lexicon.

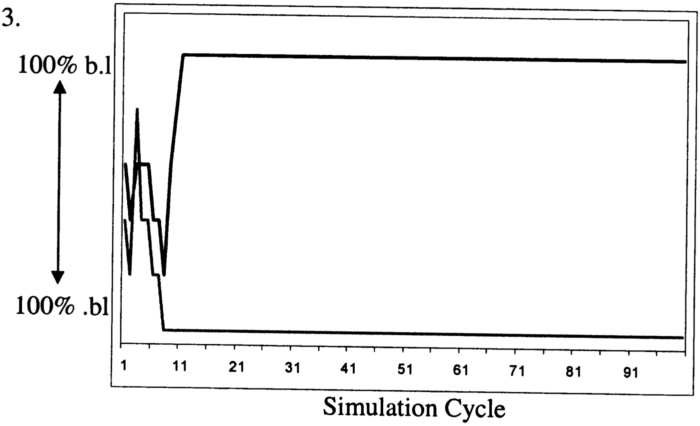
Two different runs with this lexicon are shown in Figure 3. The ordinate represents the syllabification of production targets from the lexical entry labeled *ubli*, where the top of the scale represents 100% [ub.li] targets, and the bottom

---

Wedel (forthcoming) for a version in which lexical entries can freely evolve.

represents %100 [u.bli] targets. The number of rounds is given on the abscissa. Because in the initial lexicon, all lexical entries are seeded with exemplars with opposite syllabifications, the outputs of the initial round should cluster around 50% [ub.li], [u.bli], as can be seen to be the case in both runs. However, as suggested above, any departure from 50% within any lexical entry should rapidly push them to settle on one syllabification or the other. All of the lexical entries share at least some features with one another, however, with the result that many implications derived from the lexicon will apply to many or all of the lexical entries, leading the entire lexicon to eventually veer toward a common syllabification as different syllabifications within lexical entries compete with one

Figure 3.



another. This can be seen in both runs of the simulation in Figure 3, where a global syllabification for the entire lexicon is rapidly reached, even if it differed from an early trend within the *ubli* entry. (The behavior of other lexical entries are not shown here, as they all reach a consensus syllabification at approximately the same time.)

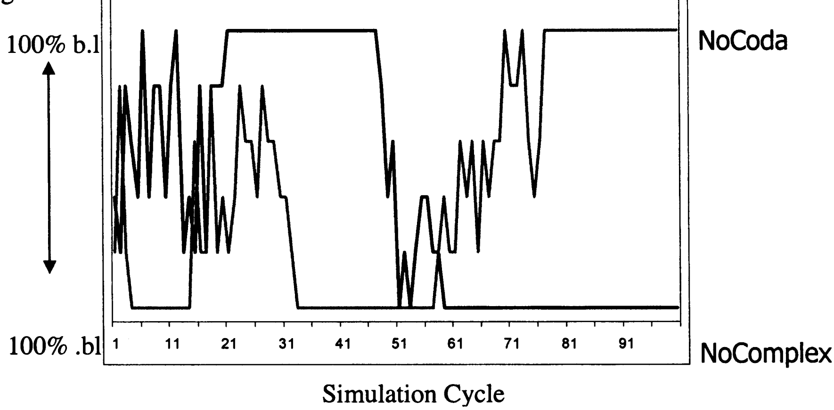
Note that the two runs shown converge on opposite syllabifications. Although there is no way to predict ahead of time which syllabification will sweep the lexicon, because the system is analogically structured, we *can* predict that sooner or later one syllabification will indeed win out. This is precisely analogous to the cellular automata simulation shown in Figure 1, which though begun with equal numbers of light and dark squares, will always end up uniformly one or the other color. This simulation then, was simply a fancier way of showing again that categorical behavior is always a basin of attraction in analogical systems.

8. Addition of external noise results in oscillation between extremes

The following simulation is run with the same lexicon shown in Figure 3, but with addition of two biases in the performance filter, one against codas, and the other against complex onsets. The first changes any word internal coda into an onset, at a rate of 10%, while the other changes any word internal onset followed by a consonant into a coda, again at a rate of 10%. All the possible production targets that this lexicon can produce are therefore going to violate one of these biases or the other, leading to the addition of balanced, but stochastic noise in performance. Results of two runs of this simulation are shown in Figure 4, where again, the lines represent the percentages of the performance targets of the two *ubli* lexical entries that have one or the other syllabification, thereby violating one or the other bias.

The behavior of the system is similar to that without biases in performance, except that instead of getting locked into one syllabification, the lexicon now

Figure 4.

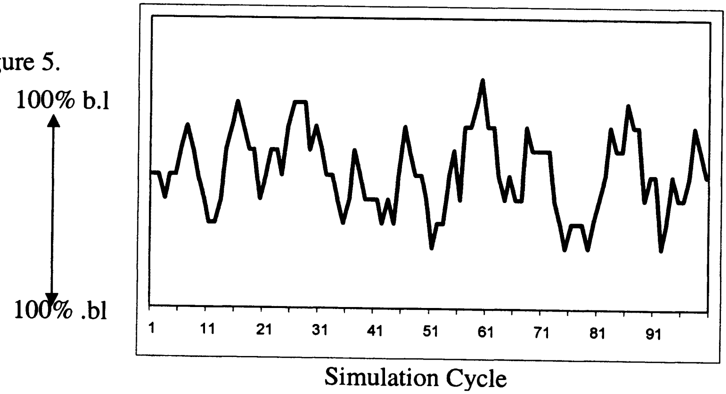


oscillates between categorical extremes in syllabification. The ability to emerge from a state in which all exemplars uniformly display one or the other syllabification derives from the biases we added into performance, which periodically alter production targets by changing their syllabification.

9. Categorical behavior is dependent on analogical structure

To show that categorical behavior is in fact crucially dependent on analogical pressure in this system, the next simulation is carried out with all ties between and within lexical entries severed: each lexical entry reproduces itself on the basis of one associated exemplar, with no reference to what any other lexical entry has done. The results, shown in Figure 5, show precisely what we expect in the presence of evenly matched, contradictory biases: syllabification behavior that simply oscillates around the mean.

Figure 5.



10. The development of patterns conforming to strict domination

Optimality Theory's restrictiveness lies in its claim that there is a limited set of universal constraints, and that there is a limited mechanism for their interaction, i.e. that the choice of optimal outputs proceeds through satisfaction of constraints in strictly ranked order. The principle of strict domination, in particular, specifies that ranking is absolute: no degree of potential violation of lower ranked constraints can ever compel violation of a higher ranked constraint. These limitations allow OT to predict that certain patterns cannot exist.

However, while strict domination allows OT to accurately describe many phonological systems, it sits uneasily with the notion, held by many, that constraints in OT are directly or indirectly related to physical factors external to the grammar. This unease arises because it is difficult to see how physiological constraints on articulation or perception would not interact additively in some overall performance cost.

The failure of grammars to reflect many of the possible levels of markedness interaction can be restated as a failure of grammatical patterns to reflect the fine-grained distinctions in difficulty that must exist (Gordon 2002). We saw above that when we model the effect of two opposing biases in an analogically structured lexicon, categorical behavior emerges from gradience. In the following sections, we will see that when multiple interacting constraints are modeled, similar categoricity evolves as well, producing grammatical behavior consonant with the strict domination principle of OT.

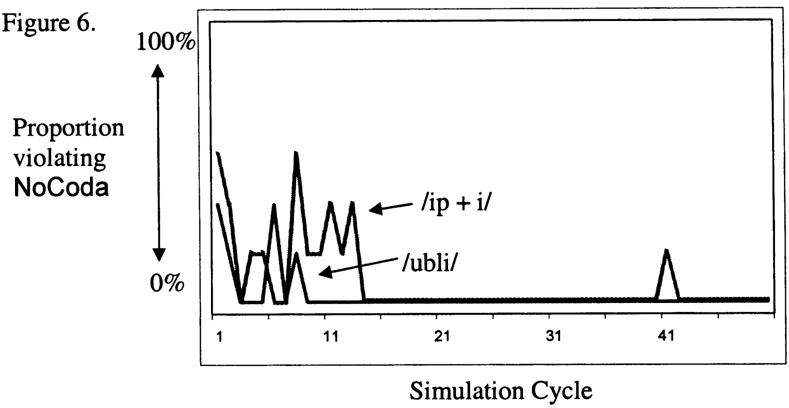
11. Setting up multiple constraint conflict in the simulation

The simulations described in this section are based on a lexicon with three classes of lexical entries, of the shapes VCCV, VC + V and VC + CV, where '+' represents a morpheme boundary. Lexical entries in the latter two classes can be thought of as comprising a stem followed by a suffix. For expository ease, I'll refer to these classes by the class members, 'ubli', 'ip + i' and 'ip + ra' respectively. The lexicon has 8 entries in each class with 5 exemplars each.

The performance filter is outfitted with three biases, NoCoda and NoComplex, and a third, abbreviated AlignStem, which operates to move syllable boundaries to coincide with stem boundaries. For the 'ubli' class, the two syllabifications trigger NoCoda or NoComplex, while for the 'ip + i' class, the two syllabifications trigger NoCoda or AlignStem, respectively. Crucially, note that for the 'ip + ra' class, one possible syllabification triggers both the NoComplex and the AlignStem biases, while the alternative triggers only the NoCoda bias. This situation gives us a chance to look test the simulation for its ability to reproduce strict domination patterns. Since for the 'ip + ra' word class, both NoComplex and AlignStem are triggered by the same syllabification, these biases both contribute to the total 'badness' of that syllabification. Hence, these biases should jointly contribute to the pressure to grammaticalize the alternative, NoCoda violating syllabification.

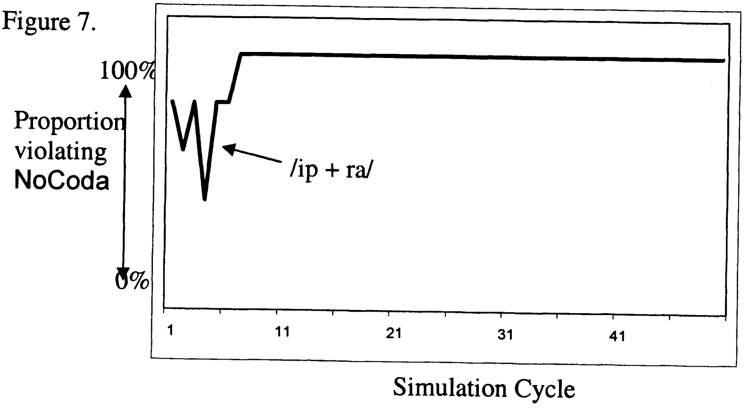
To give these biases an opportunity to do so, the simulation will operate with the bias strengths set such that NoCoda (15%) is a stronger bias than either NoComplex (10%) or AlignStem (10%) alone, but that the latter two together outweigh NoCoda. To show that these relative bias strengths do in fact result in the expected biases in output form, I show typical results of a simulation run over a lexicon containing just the /ubli/ and /ip + i/ word classes in figure 6, and a simulation run over a lexicon containing the /ip +ra/ class in isolation in figure 7.

Figure 6.



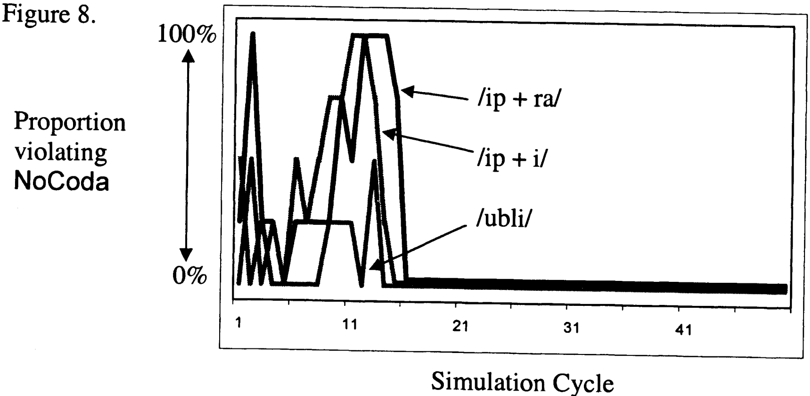
So far then, we see that when the lexicon contains only lexical entries of the /ubli/ and /ip + i/ classes (Fig. 6), the NoCoda pattern dominates the possible NoComplex and AlignStem patterns. On the other hand, when the lexicon contains only lexical entries of the /ip + ra/ class (Fig. 7), the joint NoComplex/AlignStem pattern dominates the possible NoCoda pattern. A representative simulation with a lexicon containing all three classes *together* is shown in figure 8. As before, the /ubli/ and /ip +i/ classes evolve to categorically

Figure 7.



assemble production targets without codas. However, we find that after some hesitation, the /ip + ra/ class does likewise, contradicting its behavior when evolving alone.

Figure 8.



The crucial difference between the simulation in figures 7 and 8 lies in the fact that in the latter, the /ip + ra/ class evolves not influenced only by the net bias in performance for output forms with codas in that class, but also influenced indirectly by performance's bias for forms *without* codas in the other two classes.

12. Conclusions

The simulations results presented above support the prediction that domination of one possible phonological pattern over another will be a structurally defined basin of attraction in a system in which production targets are constructed by reference to multiple forms at multiple levels of structure. In such a system, reversion to the mean sets up positive feedback loops over production and perception with the result that categorical dominance, both relative to pairs of biases, and to larger interacting bias sets, is a spontaneously reached, stable evolutionary state. Neither domination nor strict constraint domination need to be stipulated in the model, as they are straightforward consequences of lexicon structure and the mechanism by which production targets are assembled.

References

Abler, W. L. 1989. On the particulate principle of self-diversifying systems. *J. Social Biol.* Struct.12:1-13.

Blevins. J. To appear. *Evolutionary Phonology: The emergence of sound patterns.* Cambridge University Press.

Chomsky, N. and Halle, M. 1968. *The Sound Pattern of English.* New York: Harper and Row.

Goldinger, S. D. (1996) Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning Memory and Cognition* 22: 1166-1182.

Goldinger, S. D. (2000) The role of perceptual episodes in lexical processing. In Cutler, A., McQueen, J. M. and Zondervan, R. Proceedings of SWAP (Spoken Word Access Processes), Nijmegen, Max Planck Institute for Psycholinguistics. 144-159.

Gordon, M. 2002. A phonetically-driven account of syllable weight. *Language* 78:51-80.

Hintzman, D. L. 1986. "Schema Abstraction" in a Multiple-Trace Memory Model. *Psychlogical Review* 93:411-428.

Johnson, K. 1997. Speech perception without speaker normalization. In Johnson, K. and Mullennix J. W. (eds.) Talker Variability in Speech Processing. San Diego: Academic Press.

Nicolas G. and Prigogine I. 1977. *Self-organization in non-equilibrium systems: From dissipative structures to order through fluctuations.* New York: Wiley.

Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Bybee, J and P. Hopper (Eds.) *Frequency effects and the emergence of linguistic structure.* John Benjamins, Amsterdam, 137-157.

Pierrehumbert, J. 2002. Word-specific phonetics. In Gussenhoven, C and Warner, N. (Eds.) *Laboratory Phonology* 7. Berlin; New York : Mouton de Gruyter.

Prince, A. and Smolensky, P. 1993. *Optimality Theory: Constraint interaction in generative grammar.* ms. University of Colorado at Boulder CO and Rutgers University, New Brunswich, NJ.

Tenpenny, P. L. 1995. Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin and Review* 2:339-363.

Wedel, A. (forthcoming). Self-Organization and the development of higher-order phonological patterns. Ph.D Thesis, University of California, Santa Cruz, CA.

Department of Linguistics
University of Arizona
1100 E. University Blvd.
Tucson AZ 85721

Some Methodological Issues in Phonetic Typology Research: Cantonese Contour Tone Revisited

ALAN C. L. YU
University of California, Berkeley and McGill University

0. Introduction

Gordon (1999, 2001) argues that the distribution of Cantonese contour tones is typologically unusual: CVO and CVVO syllables (O = obstruents) do not host contour tones while phonemic CV syllables do. This pattern violates the generalization that longer syllables should be able to host contour tones if short syllables do in the language. Gordon rationalizes the Cantonese pattern with an acoustic analysis of the relevant syllables: rime duration of CV syllables is significantly longer than that of CVO and CVVO syllables. Gordon argues that this finding brings Cantonese into conformity with his implicational hierarchy of tone bearing ability, whereby "the tolerance of contour tones on syllables which are inherently less well suited to carrying tonal information implies the tolerance of contour tones on syllables which are better suited to manifesting tone" (Gordon 2001: 447).

This study addresses two problems with Gordon's study: one theoretical and one empirical. From a theoretical perspective, while I generally agree with Gordon's conclusion that certain syllable types are more suitable for manifesting contour tones than others (see also Zhang 2001 for a similar proposal), I find his interpretation of the Cantonese facts questionable. Gordon assumes that the non-occurrence of lexical rising tones in checked syllables (i.e. CVO and CVVO syllables) is a principled gap requiring a synchronic explanation. However, as I will demonstrate, it is in fact an accidental gap resulting from historical developments in Cantonese, and thus does not require a synchronic explanation. The details of this argument appear in section 2.

On the empirical side, Gordon's claim that no contour tones exist on CVO and CVVO syllables is simply false. As Section 3 illustrates, contour tones occur on CVO and CVVO syllables on two robust morphological contexts. Gordon, while acknowledging this in a footnote, dismisses them as irrelevant since such syllables are derived rather than lexical. It is, however, unclear why a phonetic study should *a priori* rule out morphologically derived data. Following this practice in general would limit phonologists to seeking phonetic explanations only for tautomorphemic phonological structures, rather than to surface patterns of words or phrases, surely an untenable position.

Although seriously flawed in the two respects just mentioned, Gordon's essential claim does still find support in Cantonese, in an area where he wasn't looking. In section 4 I offer phonetic evidence from Cantonese that supports Gordon's claim that contour tones prefer to occur in longer syllables than shorter ones. This study examines the phonetic realization of morphologically- and sandhi-derived rising toned CVVO syllables in Cantonese, revealing that these syllables are significantly longer than their level tone counterparts.

1. Typology without diachrony

Cantonese has six tone classes. Traditional Chinese philology treats syllables with final stops (i.e. checked syllables) as distinct tone classes (i.e. checked tones), which yields nine tones instead. Until recently, there was also a contrast between high level and high falling. However, this distinction has collapsed for most speakers today. Open syllables are generally treated as phonemically CV since there is no phonemic vowel length contrast in Cantonese, even though phonetically they are more accurately represented as CVV.

(1)	Tone	Examples
	55 (~53)	si 'poetry'
	11 (21)	si 'time'
	35	si 'to send'
	13 (23)	si 'market'
	33	si 'to try'
	22 (21)	si 'affairs'
	55	sɪk˺ 'to know'
	33	sɛːk˺ 'to kiss'
	22	sɪk˺ 'to eat'

As noted above, Gordon contends that the distribution of Cantonese contour tones is typologically unusual since checked syllables do not host contour tones while phonemic CV syllables do. One might rephrase this observation in terms of the question: why do no checked syllables in Cantonese host rising tones, whereas both rising and level tones occur in CV syllables? Gordon, seeking a synchronic, phonetically motivated answer, attributes the distribution of rising tone to the shorter phonetic duration of checked syllables relative to CV syllables. The implicit assumption behind this is that rising tones, as a type of contrastive tones in Cantonese, should be evenly distributed among the different syllable types, *all things being equal*. The crucial part of Gordon's assumption is, of course, the last four words – all things being equal. If there were evidence, for example, that the grammar of Cantonese flattens lexical contour tones when they would otherwise fall on checked syllables, such an account would be motivated. However, there is no evidence for such a process, either synchronically or diachronically, in Cantonese; thus there is no direct evidence that the absence of rising tone on checked syllables in Cantonese is causally related to the shorter duration of these syllables. In what follows, I will demonstrate that all things were indeed ***not*** equal in the history of Cantonese tonal development. Smooth syllables that host rising tones have quite a different pedigree from that of the checked syllables, making Gordon's synchronic phonetic account anomalous.

1.1 From Middle Chinese to Cantonese

Middle Chinese (*circa.* AD 200 to AD 900; MC), reflected in the pronouncing dictionary *Qieyun* (AD 601), has four tonal categories:

(2)

	Traditional name	gloss
I	ping	'level'
II	shang	'rising'
III	qu	'departing'
IV	ru	'entering'

Tones I-III were associated with smooth syllables (i.e. CV, CVV, CVN), while Tone IV occurs exclusively with checked syllables (i.e. CVO & CVVO). Phonetically speaking, Ting (1996:152), for example, reconstructs the following tonal value for MC:

(3)

	Traditional name	reconstructed phonetic values
I	Level 'ping'	level, probably low
II	Rising 'shang'	high-rising
III	Departing 'qu'	falling, probably mid-falling
IV	Entering 'ru'	abrupt and short

The loss of the initial voicing contrast gave rise to the high versus low register distinction in Cantonese tones for all four MC tones. The high register checked tones further bifurcated into high and mid level tones. The correspondence between MC tonal categories and modern Cantonese tones are given in (3). The notation *a/b* after the MC tonal category corresponds to the two tonal registers after the loss of the initial voicing contrast in obstruents.

(4)

Middle Chinese	Register Split	Cantonese
Level 'ping'	Ia	55
	Ib	11 (21)
Rising 'shang'	IIa	35
	IIb	13 (23)
Departing 'qu'	IIIa	33
	IIIb	22 (21)
Entering 'ru'	IVa	55

IVa	33
IVb	22

The history of Cantonese tones, thus, clearly demonstrates that the two types of rising tones in the language (i.e. 35 and 13) originated from the same source, namely, MC tone II. This also speaks to the fact that the reason why checked syllables do not host contour tone has little to do with its intrinsic shortness. CV syllables tolerate rising tones while checked syllables do not because the development of rising tones in Cantonese was restricted to the set of MC tone II syllables, which are all smooth syllables. The source of the contour nature of the MC tone II and III is a controversial matter. Haudricourt (1954a, b, 1961) hypothesized that MC tone II and III originated from *CVʔ* and *CVs* respectively. If this theory is correct, then the fact that checked syllables in Cantonese do not generally host contour tones is presumably due to the fact that these syllables have not lost their final consonant.[1] Thus, contour tones were unable to originate in checked syllables – but for reasons completely unrelated to phonetic duration.

To be fair, Gordon's methodological assumptions toward phonetic typological studies are not unique. Thus, the problem illustrated by this case study is actually symptomatic of the general methodology of phonetic typology. The results of a synchronic typological survey only allow us to establish inductive generalizations. The actual locus of explanation, however, lies in the diachrony, that is, in the convergence of historical developments. By ignoring this fundamental aspect of typological research, phonetic typologists risk obscuring the actual insights behind their own observations.

Given this understanding of the distribution of rising tones in Cantonese, does this mean that Cantonese is irrelevant to the discussion of the general properties of contour tone distribution? The answer here is undoubtedly no. In the second half of this paper, I look at another problem in Gordon's study of Cantonese contour tone distribution, namely, Gordon's claim that checked syllables cannot host contour tones in Cantonese.

2. Rising tones in checked syllables in Cantonese

Despite the standard description, rising tone in CVVO syllables is indeed found in Cantonese as the result of two distinct, characteristically colloquial, processes in the language: the so-called 'changed tone' and an elision-induced sandhi process.

2.1 Changed tone in Cantonese

A historical, possibly marginally productive synchronic, process, known as *changed tones*, derived mid-rising tones from semantically related level toned syllables (5).

[1] Thanks to Larry Hyman for pointing this out to me.

(5)

Level tone	Gloss	Rising tone	Gloss
ts^{h}at˺33	‘to brush’	ts^{h}at˺35	‘a brush’
p^{h}ak˺33	‘to pat’	p^{h}ak˺35	‘a racket’
kɛp˺33	‘to clip’	kɛp˺35	‘a clip’
t^{h}ɔk˺33	‘to support’	t^{h}ɔk˺35	‘a stand’
kak˺33	‘to separate’	kak˺35	‘a square’
tʃ^{h}ap˺33	‘to insert’	tʃ^{h}ap˺35	‘an insert’
tsɔk˺22	‘to chisel’	tsɔk˺35	‘a chisel’
tip˺11	‘to pile up’	tip˺35	‘a plate’

The productivity of the changed tone process in unclear. Its function is generally nominalizing, but with various complex semantic nuances (Jurafsky 1988). As Kam (1977) reported, many native speakers do not recognize any relationship between the derived forms and its alleged base. The situation is perhaps analogous to the relationship between *brother* and *brethren* in English. However, it should be noted that the number of these changed tone derived syllables is sizable.

2.2 Sandhi derived rising tones in Cantonese

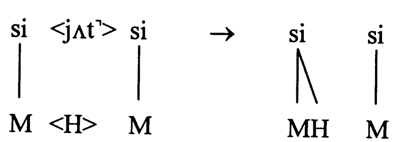
Sandhi can also give rise to rising tones in Cantonese. The verbal diminutive reduplication construction in Cantonese is formed by the reduplication of a verb root; the base and its reduplicative copy are separated by the word *jʌt˺55* ‘one’. This medial high-level tone word, however, is often omitted in conversational speech style, creating a sandhi situation. The elision of this medial syllable (i.e. *jʌt˺55* ‘one’) causes the first syllable to change to a mid-rising toned syllable (6).

(6)

si33 jʌt˺55 si33	→	si35 si33	‘to give it a try’ (lit. try one try)
p^{h}ɑt˺ 33 jʌt˺55 p^{h}ɑt˺ 33	→	p^{h}ɑt˺ 35 p^{h}ɑt˺ 33	‘to hit a little’ (lit. hit one hit)
ts^{h}ɑɑt˺33 jʌt55 ts^{h}ɑɑt˺33	→	ts^{h}ɑɑt˺35 ts^{h}ɑɑt˺33	‘to brush a little’ (lit. brush one brush)

Phonologically, this sandhi process is often characterized as the re-association of a high tone, set afloat by the deletion of an originally high toned syllable, with the preceding level toned syllable (Chen 2000: 60).

(7)



This is a productive construction, meaning that rising tones in checked syllables are quite commonly derived in Cantonese.

To summarize briefly, checked syllables in Cantonese can in fact host rising tones, albeit only through morphological and post-lexical processes such as Changed Tone and sandhi respectively. It is therefore extremely puzzling that Gordon chooses to ignore these forms in his phonetic study of Cantonese tone and syllable type. One should note that since his study focuses on the phonemic status of syllable types, morphological and syntactic considerations should therefore be irrelevant to begin with.[2]

3. Contour tone-induced lengthening

In the remainder of this paper, I present a phonetic study of these checked syllables with derived rising tones in Cantonese, showing that syllables with deriving rising tones are significantly longer than their level toned counterparts. This novel finding, while based on words which counterexemplify Gordon's attempt to derive Cantonese tonal contrasts within roots from phonetic principles, does lend support for Gordon's more general claim that durationally longer syllables are better host to rising tones than shorter ones.

3.1 Methodology

Two native speakers of Cantonese (a college-age male and a middle-age female) recited a list of Cantonese target disyllabic words/phrases in the carrier phrase /ŋɔ tɔk˺ ___ pɛi nɛi tʰæŋ/. Two sets of tokens were recorded. In Set 1, the target words were eleven doublets of CVVO Cantonese syllables. The doublets consisted of segmentally identical CVVO syllables with lexical level tones and their morphologically derived rising toned counterparts (8). While the derived rising tones are all mid-rising (˧˥ or 35), the lexical level tones may be extra-low (˩ or 11), low (˨ or 22), or mid (˧ or 33) (see Appendix for the entire corpus).

(8) Lexical level ˧ (or 33)
(tɔɔ 55) **tsʰɐt˺33** 'to brush (more)'

Morphologically derived ˧˥ (or 35)
(tsʊk˺55) **tsʰɐt˺35** 'a (bamboo) brush'

---

[2] On a related methodological note, it is interesting to point out that three out of five of the tokens chosen in Gordon's phonetic study are based on reading pronunciations of standard Chinese characters (e.g., *mà* 'to curse', *làp* 'to stand', *sàm* 'very'). While the phonetic differences between reading versus colloquial pronunciations have not been extensively studied, some difference has been noted in the literature (e.g., Bauer & Benedict 1997:94-107).

In Set 2, the target syllables are couched within the sandhi context described in section 2.2. The target syllables consisted of eight CVVO Cantonese syllables reduplicated according to rules of the truncated diminutive reduplication construction (9).

(9) Sandhi-derived 35
ts^hɐt˺35 ts^hɐt˺ 33 'to brush a little'

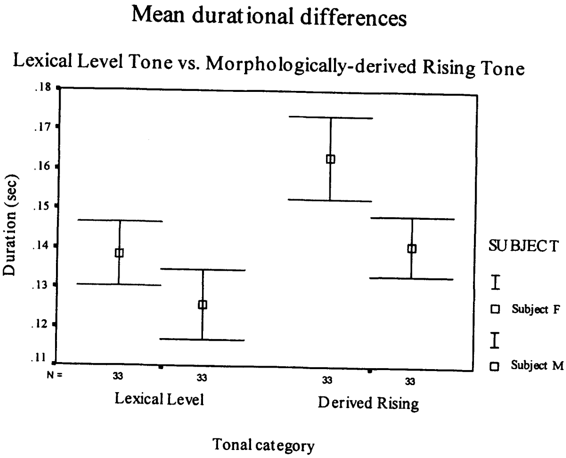
3.2 Analysis and results

The duration of the target syllables in Set 1 were measured. Table 1 shows the mean durations of the CVVO syllables with lexical level or morphologically derived tones for both subjects. A one-way ANOVA reveals that the duration of the CVVO syllables with the morphologically derived rising tone is significantly longer than their lexical level toned counterparts ($p < 0.05$) for both subjects (see Figure 1).

Table 1: Mean durations of CVVO syllables with lexical level tones and morphological-derived rising tones.

	Subject F	Subject M
Lexical Level	0.1384 (0.0229) N=33	0.1255 (0.0251) N=33
Morphologically-derived Rising	0.1631 (0.0293) N=33	0.1409 (0.0217) N=33
	(p = 0.000)	(p= 0.010)

Figure 1: Mean durations of CVVO syllables with lexical level tones and morphological-derived rising tones.

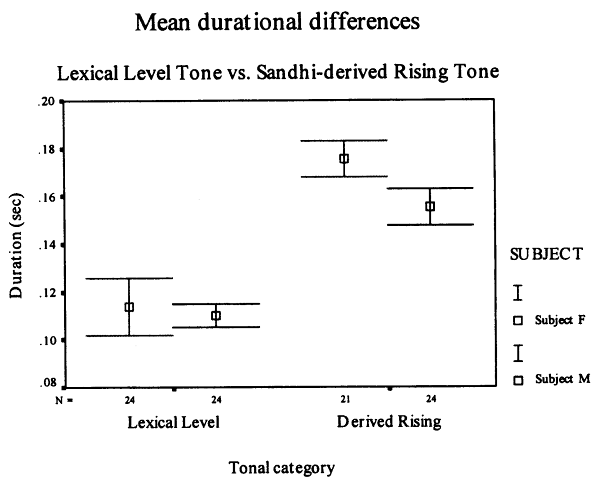


To control for any speaking rate effects, syllables with sandhi-derived rising tones (i.e. the Set 2 tokens) are compared only to the syllable immediately following it. This syllable is identical to the target CVVO except that it preserves its original level tone (e.g., *tsʰɐt˧˥35 tsʰɐt˧33*). Table 2 shows, for both subjects, the mean durations of CVVO syllables with lexical level or sandhi-derived tones. A one-way ANOVA reveals that the sandhi-derived rising toned syllables are significantly longer than their level toned counterparts ($p < 0.05$) for both subjects (see Figure 2).

Table 2: Mean durations of CVVO syllables with lexical level tones and sandhi-derived rising tones

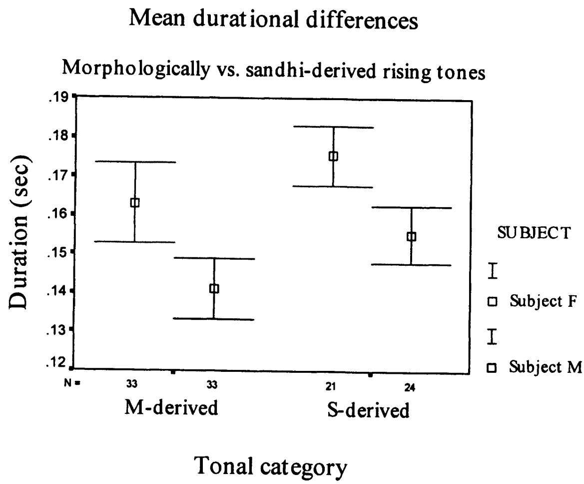
	Subject F	Subject M
Lexical level	0.1138 (0.0282) N=24	0.1100 (0.0121) N=21
Sandhi-derived Rising	0.1809 (0.0241) N=24	0.1552 (0.0176) N=24
	($p = 0.000$)	($p = 0.000$)

Figure 2: Mean durations of CVVO syllables with lexical level tones and sandhi-derived rising tones



The mean durational differences between the CVVO syllables with morphologically-derived and sandhi-derived rising tones are compared in Figure 3. A one-way ANOVA shows that, for Subject M, the sandhi-derived rising toned syllables are significantly longer than the morphologically-derived ones ($p < 0.05$), while there is no significant durational difference in the speech of Subject F between these two types of derived rising toned syllables (see Figure 3).

Figure 3: Mean durations of CVVO syllables with morphological (M)- and sandhi (S)-derived rising tones



4. Discussions

This study demonstrates that syllables with derived contour tone in Cantonese are significantly longer than their underived level toned counterparts. This phenomenon in Cantonese is, to my knowledge, unheard of prior to this study. The results of this study also support the phonetic explanation of the durational properties of contour tone bearing units and Gordon's implicational hierarchy. *A priori*, the lengthening effect is unexpected. By comparing segmentally identical syllables in their derived versus underived contexts, I was able to ascertain that contour tones indeed prefer a longer duration of the sonorous phase of its host.

Interestingly, it should also be noted that, while it is commonplace for rising tones to have a more limited distribution than the other tone types, the process of rising tone-induced lengthening appears to be relatively rare across the world's languages. Zhang (2001), for example, in his typological survey of the distribution and the phonological properties of contour tones, reports only three documented cases (i.e. Mitla Zapotec (Briggs 1961), Wuyi Chinese (Fu 1984), Gã (Paster 1999)). He, however, attributes this asymmetry to greater attention devoted to documenting the restrictions of contour tones on syllable types in his data sources. The durational change of contour tone syllables might have escaped the attention of many grammarians. Thus, instrumental studies demonstrating the effect of contour tone-induced lengthening such as the one presented here should be a welcoming development.

As mentioned in the last section, the duration of the sandhi-derived rising toned syllables is significantly longer than the duration of the morphologically-derived ones in Subject M's speech. While the difference between these two types of derived syllables in Subject F's speech is not statistically significant, the overall trend, nonetheless, mimics that of Subject M. That is, Subject F's sandhi-derived rising toned syllables are generally longer than the morphologically-derived ones. *A priori*, this is unexpected, since both types of derived syllables host a type-similar rising tone, namely a 35 mid-rising tone. Several factors, however, might contribute to this disparity. First, the contexts which these two types of syllables appear are different. Following the standard metrical analysis of Cantonese phonology, the target syllables in the sandhi environment occurs in the head of a trochaic foot, while the target syllables in the derived morphological context is in the dependent position of a trochaic foot. Thus, the extra duration of the target syllables in the sandhi environment might be the result of some form of a stress-induced lengthening effect. However, since no available study has looked at the phonetic effects of stress in Cantonese, this explanation remains speculative. The target syllables in sandhi-derived environment might also be longer due to some form of a compensatory lengthening effect as the result of the elision of the syllable immediately following the target syllable. Further study is underway to discern the actual mechanism(s) contributing to this durational disparity.

5. Conclusions

This paper has focused on Gordon's analysis of the distribution of contour tones in Cantonese, showing that the absence of lexical contour tones in checked syllables is better explained through diachronic means rather than synchronic phonetic properties. While durational factors may correlate with certain tonal patterns, the actual direction of influence must be established on an individual basis. At least in Cantonese, durational variation is a side effect of tonal realization, rather than a predictor of it. I also point out that contour tones indeed occur in checked syllables in Cantonese, but only in derived contexts. By ignoring these syllables in his phonetic study, Gordon misses an important generalization about Cantonese: contour tones induce a lengthening effect on its host syllable. A study looking at the durational properties of these derived contour toned syllables in Cantonese confirms the predictions of the phonetic interpretation of tone bearing ability. That is, the derived contour toned syllables are significantly longer than their lexical level toned counterparts. The use of contrasts introduced by morphological and sandhi considerations in phonetic studies, I argue, provide a powerful tool to testing hypotheses in phonological and phonetic theories alike.

APPENDIX: Cantonese corpus

(10) Lexical level tone target syllables (The target syllable is underlined)

Word	Gloss	Word	Gloss
sa:55 kɔk˺33	(a place name)	kʌp˺55 tʃʰap˺33	'to insert quickly'
tɔ:55 tsʰat˺33	'to brush more'	kʌp˺55 tsɔk˺22	'to chisel quickly'
tɔ:55 pʰak˺33	'to pat more'	fʊŋ55 jip˺22	'maple leaf'
hæŋ55 tʰɔk˺33	'to support a little'	kwɔŋ55 mɔk˺22	'a screen'
fʌn55 kak˺33	'to separate squarely'	hak˺55 mʌk˺22	'black ink'
		hæŋ55 tip˺11	'to pile up gently'

(11) Morphologically-derived mid-rising target syllables

Word	Gloss	Word	Gloss
sa:55 kɔk˺35	(a type of food)	kʌʌm55 tʃʰap˺35	'a golden insert'
tsʊk˺55 tsʰat˺35	'a bamboo brush'	kʌʌm55 tsɔk˺35	'a golden chisel'
pɔ:55 pʰak˺35	'a ball racket'	kɛi55 jip˺35	'a propeller'
pui55 tʰɔk˺35	'a cup stand'	kɔn55 mɔk˺35	'a liver'
fɔŋ55 kak˺35	'a square'	hak˺55 mʌk˺35	'a mole'
		kwɔŋ55 tip˺35	'a CD'

(12) Sandhi-derived mid-rising target syllables

Word	Gloss
tsʰat˺35 tsʰat˺33	'to brush a little'
pʰak˺35 pʰak˺33	'to hit a little'
kɛp˺35 kɛp˺33	'to clip a little'
tʰɔk˺35 tʰɔk˺33	'to support a little'
kak˺35 kak˺33	'to separate a little'
tʃʰap˺35 tʃʰap˺33	'to insert a little'
tsɔk˺35 tsɔk˺22	'to chisel a little'
tip˺35 tip˺22	'to pile up a little'

References

Bauer, Robert and Paul, Benedict. 1997. *Modern Cantonese phonology*. New York: Mouton.

Briggs, Elinor. 1961. *Mitla Zapotec grammar*. Mexico: Instituto Lingusitico de Verano.

Chen, Matthew. 2000. *Tone sandhi: pattern across Chinese dialects*. Cambridge: Cambridge University Press.

Fu, Guotong. 1984. Wuyi fanyan de liandu biandiao [Tone sandhi in the Wuyi dialect]. *Fangyan* 1984(2): 109-127.

Gordon, Matthew. 1999. *Syllable Weight: Phonetics, Phonology and Typology*. PhD dissertation, University of California, Los Angeles.

Gordon, Matthew. 2001. A typology of contour tone restrictions. *Studies in Language* 25: 405-444.

Haudricourt, A. G. 1954a. De l'origine des tons en Vietnamien. *Journal Asiatique* 242: 69-82.

Haudricourt, A. G. 1954b. Comment reconstruire le chinois archaïque. *Word* 10: 351-364.

Haudricourt, A. G. 1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient. *Bulletin de la Société Linguistique de Paris* 56:163-180.

Jurafsky, Daniel. 1988. On the semantics of Cantonese changed tone: Women, matches, and Chinese broccoli. In *Proceedings of 14th annual meetings of the Berkeley Linguistics Society*. Berkeley, CA. 304-318.

Kam, Tak Him. 1977. Derivation by tone change in Cantonese: A preliminary survey. *Journal of Chinese Linguistics* 5: 186-210.

Paster, Mary. 1999. *Issues in the tonology of Gã*. Senior honor thesis, The Ohio State University.

Ting, Pang-Hsin. 1981. Tonal evolution and tonal reconstruction. In C.T. James Huang and Y. H. Audrey Li (eds.), *New Horizons in Chinese Linguistics*. Kluwer Academic Publishers. 141-159.

Zhang, Jie. 2001. *The Effects of Duration and Sonority on Contour Tone Distribution--Typological Survey and Formal Analysis*. PhD dissertation, University of California, Los Angeles.

Alan C. L. Yu
Department of Linguistics
1203 Dwinelle Hall
University of California
Berkeley, CA 94720-2650

aclyu@socrates.berkeley.edu

BERKELEY LINGUISTICS SOCIETY

1203 Dwinelle Hall, University of California, Berkeley, CA 94720-2650 USA
Telephone and Fax: (510) 642-5808; e-mail: bls@socrates.berkeley.edu
http://www.linguistics.berkeley.edu/BLS

Book Prices

		Parasession & Special Session (S) Topics	Individuals	Institutions
BLS-1	(1975)		$10.00	$15.00
BLS-2	(1976)		$10.00	$15.00
BLS-3	(1977)		$10.00	$15.00
BLS-4	(1978)		$10.00	$15.00
BLS-5	(1979)		$10.00	$15.00
BLS-6	(1980)		$10.00	$15.00
BLS-7	(1981)		$10.00	$15.00
BLS-8	(1982)		$14.00	$20.00
BLS-9	(1983)		$10.00	$15.00
BLS-10	(1984)	Subordination	$14.00	$20.00
BLS-11	(1985)	Poetics, Metrics, and Prosody	$14.00	$20.00
BLS-12	(1986)	Semantic Typology	$14.00	$20.00
BLS-13	(1987)	Grammar and Cognition	$14.00	$20.00
BLS-14	(1988)	Grammaticalization	$14.00	$20.00
BLS-15	(1989)	Theoretical Issues in Language Reconstruction	$14.00	$20.00
BLS-16	(1990)	Legacy of Grice	$16.00	$20.00
BLS-16S	(1990)	General Topics in American Indian Linguistics	$8.00	$12.00
Author Index to BLS 1-16S			$2.00	$4.00
BLS-17	(1991)	The Grammar of Event Structure	$16.00	$20.00
BLS-17S	(1991)	African Language Structures	$8.00	$12.00
BLS-18	(1992)	Place of Morphology in a Grammar	$16.00	$20.00
BLS-18S	(1992)	Typology of Tone Languages	$8.00	$12.00
BLS-19	(1993)	Semantic Typology and Universals	$16.00	$20.00
BLS-19S	(1993)	Syntactic Issues in Native American Languages	$8.00	$12.00
BLS-20	(1994)	Dedicated to the Contributions of Charles J. Fillmore	$16.00	$20.00
BLS-20S	(1994)	Historical Issues in African Linguistics	$8.00	$12.00
BLS-21	(1995)	Historical Issues in Sociolinguistics/ Social Issues in Historical Linguistics	$16.00	$20.00
BLS-21S	(1995)	Discourse in Southeast Asian Languages	$8.00	$12.00
BLS-22	(1996)	The Role of Learnability in Grammatical Theory	$16.00	$20.00
BLS-22S	(1996)	Historical Issues in Native American Languages	$4.00	$12.00
BLS-23	(1997)	Pragmatics and Grammatical Structure	$16.00	$20.00
BLS-23S	(1997)	Syntax and Semantics in Africa	$8.00	$12.00
BLS-24	(1998)	Phonetics and Phonological Universals	$20.00	$25.00
BLS-24S	(1998)	Indo-European Subgrouping and Internal Relations	$4.00	$8.00
BLS-25	(1999)	Loan Word Phenomena	$20.00	$25.00
BLS-25S	(1999)	Caucasian, Dravidian and Turkic Linguistics	$8.00	$12.00
BLS-26	(2000)	Aspect	$20.00	$25.00
BLS-26S	(2000)	Syntax and Semantics of the Indigenous Languages of the Americas	$10.00	$15.00
BLS-27	(2001)	Language and Gesture	TBA	TBA
BLS-27S	(2001)	Afroasiatic Linguistics	$10.00	$15.00
BLS-28	(2002)	Field Linguistics	$20.00	$25.00
BLS-28S	(2002)	Tibeto-Burman/Southeast Asian Linguistics	TBA	TBA
BLS-29	(2003)	Phonetic Sources of Phonological Patterns	$20.00	$25.00
BLS-29S	(2003)	Minority and Diasporic Languages of Europe	$10.00	$15.00

California Sales Tax

All orders from residents of California must include sales tax according to county of residence.

Shipping and Handling (USPS ground)

	First Book	Each Additional Book
Within U.S.	$4.00	$2.00
Canada and Mexico	$7.00	$3.50
All other countries	$9.00	$4.00

ORDERS MUST BE ACCOMPANIED BY PAYMENT.

Make checks and money orders payable to the Berkeley Linguistics Society. Checks must be in U.S. dollars and drawn on U.S. banks.

Prices and shipping costs subject to change without notice.